Principles of Electrodynamics

ALEKSEY MATVEYEV

Principles of Electrodynamics

Principles of Electrodynamics

by

ALEKSEY N. MATVEYEV

Moscow State University
Moscow, U.S.S.R.

translation editor

LEON F. LANDOVITZ

Yeshiva University
New York, N. Y.

NEW YORK

Reinhold Publishing Corporation

CHAPMAN & HALL, LTD., LONDON

Library of Congress Catalog Card Number 65-28272
Printed in the United States of America
Translated by Scripta Technica, Inc.

Translation Editor's Preface

Electromagnetic forces are probably the best understood of all the forces in nature and are vital to any thorough understanding of macroscopic processes. Thus, it is imperative that the physicist or engineer be introduced to the fundamental principles underlying our knowledge of these forces at an early stage of his education.

In this book, we have a clear, concise introduction, on the intermediate level, of all the tools necessary to handle the most important problems in electrodynamics, with emphasis on the experimental basis of significant phenomena. The book is divided into three parts: Phenomenological Electrodynamics, Electron Theory, and the Theory of Relativity. The first two parts present Maxwell's Equations and their consequences, first introducing phenomenological parameters to describe the behavior of material media and then deriving them from a more fundamental microscopic view. Einstein, through his Theory of Relativity, made possible a beautiful unification of electric and magnetic phenomena. Therefore, a discussion of the historical background which led to Einstein's theory, its fundamental concepts, and their far-reaching consequences may be found in the last part of this text. *Principles of Electrodynamics*, then, fills the need for a somewhat more advanced text on electricity and magnetism which does not assume great mathematical sophistication, but which does emphasize the basic physics.

The changes from the original Russian are minor. Mainly, they consist of either clarifications of statements which could be misinterpreted by the

unwary, or expansions of certain points which the author has not developed in detail because of the mathematical level assumed but which can be handled quite adequately by the American undergraduate.

LEON F. LANDOVITZ

Contents

Principles of Electrodynamics

PART I

Phenomenological Electrodynamics

IT IS well known that a wide range of electromagnetic phenomena may be accounted for without introducing the molecular nature of matter and the discrete nature of electric charges. In this approach, electrical and magnetic properties of a substance are described by the permittivity ϵ, the magnetic permeability μ, and the electrical conductivity λ. Charges and currents are assumed to be distributed continuously in space and are described by a charge density ρ and a current density $\mathbf{j}$. This idealized picture of matter, charges, and currents proves satisfactory in many cases. The electromagnetic theory based on this idealization is called *phenomenological electrodynamics* and is dealt with in the first part of this book.

The second part deals with the theory of electromagnetic phenomena, which takes into account the molecular structure of matter and the discrete nature of the electric charges. This is the so-called *electron theory of matter*. It provides an insight into the mechanism of a number of phenomena, which within the framework of phenomenological electrodynamics either can be described in a purely formal manner, or cannot be explained at all.

The third part of this book is devoted to the *theory of relativity*. It is not by chance that the theory of relativity

is discussed together with electrodynamics. First, this theory originally arose in connection with electrodynamics, and second, electrodynamics was, historically, the first relativistically invariant theory. Third, electrodynamics provides many examples of the application of the theory of relativity which illuminate the physical meaning of the formulas and the predictions of the theory of relativity.

Electromagnetic phenomena play an exceedingly important role in nature. At present, four types of interactions are recognized: *gravitational, electromagnetic, strong,* and *weak.* All other interactions known at present may be reduced to these four. For example, viscous forces and many others are essentially electromagnetic in origin.

The gravitational forces acting between charged particles are very small in comparison with the electrical forces between them. For example, the force of gravitational attraction between two electrons separated by a distance r is

$$F_T = G\frac{m_0^2}{r^2} \tag{1}$$

where $G = 6.7 \times 10^{-11}$ N m²/kg is the gravitational constant, and $m_0 = 9.1 \times 10^{-31}$ kg is the rest mass of the electron. On the other hand, the electrical force of repulsion acting between electrons is

$$F_e = \frac{e^2}{4\pi\epsilon_0 r^2} \tag{2}$$

where $e = 1.6 \times 10^{19}$ coulombs (coul) is the electron charge, and $\epsilon_0 = \frac{1}{4\pi} \times 9 \times 10^9$ F/m is the permittivity of empty space. From (1) and (2) it follows that

$$\frac{F_e}{F_T} = \frac{e^2}{4\pi\epsilon_0 G m_0^2} \approx 10^{43} \tag{3}$$

so that the gravitational interaction between two electrons is negligible in comparison with the electrical interaction between them. It is thus evident that the gravitational forces play practically no part in interactions between elementary particles. The gravitational forces are important only in the interaction of large neutral masses such as, for example, astronomical objects.

Specifically *nuclear* forces, which depend upon strong interactions, are at present not known exactly, but their properties have been studied in fairly great detail. It is known that these are short range forces: they are significant only when nucleons are at a distance of the order of 10^{-13} cm. At such distances, nuclear forces are very much greater than electromagnetic forces. However, they decrease rapidly as the distance increases, and soon become negligible in comparison with electromagnetic forces. It follows that nuclear forces are important in the interactions between elementary particles when the latter closely approach each other. In particular, these forces play a dominant role in the formation of atomic nuclei.

Weak interactions appear in β-decay in which fast electrons (β-particles) are emitted in the course of nuclear transformations. The decay of the neutron is an example of this process.

Of the four types of interaction known in nature at the present time, only the electromagnetic interaction can be used to control the motion of charged particles, and it is therefore exceptionally important in present day science.

One of the principal problems confronting science today is that of controlled thermonuclear reactions. At present, this problem is primarily one of electrodynamics, and reduces to the use of electromagnetic fields as "reservoirs" in which hot plasma is confined. In a way, the magnetic field is being used as a construction material.

It is recognized that ion and plasma jet engines will play an important part in future space flights, and here again electrodynamic problems are of basic importance. There is, of course, already a vast range of practical applications of electromagnetic phenomena in electrical engineering, electronics, and so forth. Electrodynamic phenomena also play a large part in astrophysics: the magnetic fields in interstellar space accelerate cosmic charged particles, solar flares are accompanied by considerable changes in the magnetic field close to the sun's surface, and the earth's magnetic field confines charged particles in the neighborhood of the earth, thus creating radiation belts. The enormous range and variety of electromagnetic phenomena ensure that their theory is an important tool in the exploration of the laws of nature.

CHAPTER 1

Maxwell's Equations as a Generalization of Experimental Fact

§1. The Electromagnetic Field. System of Units

Definitions. An *electromagnetic field* is a region of space in which electric and magnetic forces are acting. It is described by the electric field **E**, the electric induction **D**, the magnetic field **H**, and the magnetic induction **B**. These four vectors are connected by the relations†

† If the medium is inhomogeneous, the equations are correct as written except that ϵ and μ are not constant. But if the medium is anisotropic, e.g., Nicole prism, then

$$\mathbf{D}_i = \epsilon_{ij}\mathbf{E}_j$$
$$\mathbf{B}_i = \mu_{ij}\mathbf{H}_j$$

and ϵ is not a scalar but a tensor. In fact, ϵ and μ are always tensors, but for isotropic media

$$\epsilon_{ij} = \epsilon\,\delta_{ij}$$
$$\mu_{ij} = \mu\,\delta_{ij}$$

where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

and

$$Di = \epsilon Ei$$
$$Bi = \mu Hi$$

or in vector notation

$$\mathbf{D} = \epsilon\mathbf{E}$$
$$\mathbf{B} = \mu\mathbf{H}$$

$$\mathbf{D} = \epsilon\mathbf{E} \tag{1.1a}$$

$$\mathbf{B} = \mu\mathbf{H} \tag{1.1b}$$

where ϵ is the permittivity and μ the permeability.

In the absolute Gaussian system of units, which is widely used in electrodynamics, electrical quantities are measured in esu, and magnetic quantities in emu. In this system, the dimensions of **E** and **D** and **H** and **B** are the same, while ϵ and μ are dimensionless and are equal to unity in empty space. Hence, in empty space, the vectors **D** and **E** and the vectors **B** and **H** are equal, and the electromagnetic field may be described by only two vectors, **E** and **H**, which leads to a considerable mathematical simplification. However, many electrical quantities are then expressed in units which are different from the practical units (ampere, volt, coulomb, ohm, etc.). In the Gaussian system, all these quantities are measured in electrostatic units (esu) which have no individual names and are not used in practical work. It is, therefore, desirable to have a system of units which is suitable for practical work. Such a system is the international system of units, called SI (Système Internationale to distinguish it from an earlier international system), which was put forward by the International Assembly on Weights and Measures and has now been adopted by many countries, including the USSR, as the official standard. In this system, the vectors **E** and **D** have different dimensions, and ϵ is therefore a dimensional quantity which is not equal to unity in space; consequently, **E** and **D** are different even in empty space. The same may be said about **H**, **B**, and μ. Thus, in the SI system, we need the four vectors **E**, **D**, **H**, and **B** to describe the field both in material and in empty space.

The SI system of units will be used throughout this book. For electromagnetic quantities, this system is identical with the rationalized MKSA system. The dimensions and names of the fundamental units of this system are given in Appendix 2. In the SI system, ϵ and μ have the following dimensions

$$[\epsilon] = \frac{\text{amp}^2\,\text{sec}^4}{\text{Kg m}^3} = \frac{\text{coul}^2}{\text{J m}} = \frac{F}{\text{m}} \tag{1.2a}$$

$$[\mu] = \frac{\text{m Kg}}{\text{amp}^2\,\text{sec}} = \frac{\text{m Kg}}{\text{coul}^2} = \frac{H}{\text{m}} \tag{1.2b}$$

which follow from (1.1a) and (1.1b), and the dimensions of **E**, **D**, **H**, and **B**, given in Appendix 2.

In the SI system, the ampere is defined as the constant current which, flowing along two parallel straight conductors of infinite length and negligibly small circular cross section separated by a distance 1 m *in vacuo*,

gives rise to a force of 2×10^{-7} newtons per meter (N/m) between the conductors.

The force, per length l, between the two parallel conductors *in vacuo* separated by a distance a is given by

$$F = \frac{\mu_0 I_1 I_2 l}{2\pi a} \tag{1.3}$$

Using the definition of the ampere, and substituting $F = 2 \times 10^{-7}$ N, $I_1 = I_2 = 1$ amp, $l = a = 1$ m, we obtain the following value for the permeability of empty space

$$\mu_0 = 4\pi \times 10^{-7} \text{ henrys/meter} \tag{1.4}$$

To obtain the permittivity, we use the theoretical formula

$$\frac{1}{\sqrt{\epsilon_0 \mu_0}} = c \tag{1.5}$$

where c is the electrodynamic constant equal to the velocity of light in empty space ($c = 3 \times 10^8$ m/sec). Equation (1.5) is derived in Section 32. From (1.5) and (1.4) it follows that

$$\epsilon_0 = \frac{1}{4\pi\, 9 \times 10^9} \frac{F}{\text{m}} \tag{1.6}$$

Using (1.4) and (1.6), the permittivity and the permeability may also be written in the form

$$\epsilon = \epsilon' \epsilon_0 \tag{1.7a}$$

$$\mu = \mu' \mu_0 \tag{1.7b}$$

where ϵ' and μ' are dimensionless quantities, called, respectively, the *relative permittivity* and the *relative permeability*, and are numerically equal to the permittivity and permeability in the absolute Gaussian system of units.

The field vectors **E**, **D**, **H**, and **B** are, in general, functions of the coordinates and of time. The permeability μ and the permittivity ϵ are functions of the coordinates, and it is assumed that they do not depend explicitly on time.

Volume Charge Density. The distribution of charge in space is described by the *volume density* ρ, which is defined as follows

$$\rho = \lim_{\Delta V \to 0} \frac{\Delta q}{\Delta V} \tag{1.8}$$

where Δq is the charge contained in volume ΔV. This definition assumes a continuous distribution of the charge in space, and the discrete nature

of electric charge is, therefore, ignored. It is evident from (1.8) that the volume charge density is measured in coulombs per cubic meter

$$[\rho] = \text{coul/m}^3 \tag{1.9}$$

It also follows from (1.8) that the amount of charge dq contained in an element of volume dV is given by

$$dq = \rho dV \tag{1.10}$$

Current Density. The current density $\mathbf{j}$ is a vector, directed along the current at a given point. Its absolute value is

$$|\mathbf{j}| = \lim_{\Delta S \to 0} \frac{\Delta I}{\Delta S} \tag{1.11}$$

where ΔI is the current passing through an element of area ΔS which is perpendicular to the direction of the current at the given point. It is evident from (1.11) that the current density, is measured in amperes per square meter

$$[\mathbf{j}] = \text{amp/m}^2 \tag{1.12}$$

It also follows from (1.11) that the current dI passing through an element of area dS, is defined by the scalar product

$$dI = \mathbf{j} \cdot d\mathbf{S} \tag{1.13}$$

The volume charge density and the current density are functions of the coordinates and of time.

Special Features of the Electromagnetic Field Theory. The electromagnetic field is described at every point by the vectors **E**, **D**, **H**, and **B**, the values of which vary, generally speaking, with time, and also from point to point. However, this variation is not arbitrary. It takes place according to definite laws. The generation of an electric field by charges and currents also takes place according to definite laws. These laws are formulated in *Maxwell's equations*, which are satisfied by quantities which describe the electromagnetic field.

Originally, the laws governing electromagnetic phenomena were presented as relationships relating to different points in space. For example, *Coulomb's law* gives the value of the force acting between charges at different points; *Ohm's law* defines a relationship between quantities which refer to a section of a conductor, and so on. Maxwell's equations, on the other hand, formulate the laws of the electromagnetic field as relationships between values at the same point in space and at the same instant of time, and this is the characteristic feature of the mathematical description of the electromagnetic field. It follows that, in order to obtain the equations of the electromagnetic field, we must put the fundamental laws of electro-

magnetic phenomena into the form of relationships between quantities at the same point in space and at the same instant of time, i.e., we must write these laws in differential form.

§2. Differential Form of Gauss' Theorem

Field Treatment of Coulomb's Law. Coulomb's law defines the force F between two point charges e_1 and e_2 in a homogeneous medium of permittivity ϵ

$$F = \frac{1}{4\pi\epsilon}\frac{e_1e_2}{r^2} \tag{2.1}$$

where r is the distance between the charges.

From the point of view of the electromagnetic field, the interaction between the charges is as follows:

(1) A point charge, e.g., the charge e_1, creates in its neighborhood an electric field whose intensity $\mathbf{E}$ is given by

$$\mathbf{E} = \frac{1}{4\pi\epsilon}\frac{e_1}{r^2}\frac{\mathbf{r}}{r} \tag{2.2}$$

where $\mathbf{r}$ is the radius vector drawn from the point where the charge e_1 is situated to the point at which the intensity is being defined.

(2) A point charge e_2 placed in an electric field of intensity $\mathbf{E}$ experiences a force due to this field equal to

$$\mathbf{F} = e_2\mathbf{E} \tag{2.3}$$

Substituting in (2.3) the expression for $\mathbf{E}$ from (2.2), we obtain the formula for the force $\mathbf{F}$ between e_1 and e_2

$$\mathbf{F} = \frac{1}{4\pi\epsilon}\frac{e_1e_2}{r^2}\frac{\mathbf{r}}{r} \tag{2.4}$$

thus demonstrating that Coulomb's law (2.1) is contained in formulas (2.2) and (2.3).

The expression (2.3) for the force acting on a point charge placed in an electric field is quite general, and is independent of the reason for the existence of the field $\mathbf{E}$. Hence, Coulomb's law is contained in (2.2), which may be conveniently written in the form

$$\epsilon\mathbf{E} = \mathbf{D} = \frac{1}{4\pi}\frac{e}{r^2}\frac{\mathbf{r}}{r} \tag{2.5}$$

where $\mathbf{D}$ is the electric induction vector. Equation (2.5) expresses the fact that the electric induction vector of a point charge is inversely proportional to the square of the distance from the charge. Moreover, it appears from

this formula that, for a given distribution of charges, the induction vector **D** is independent of the medium; in empty space or in a medium, at the same distance from a point charge e, the electric induction vector will be the same. Formula (2.5) also indicates that the electric induction is measured in coulombs per square meter

$$[\mathrm{D}] = \mathrm{coul/m^2}$$

Gauss' Theorem. Let us work out the flux of the electric induction vector **D** through an arbitrary closed surface S surrounding a point charge e, i.e., let us find

$$N = \oint_S \mathbf{D} \cdot d\mathbf{S} \tag{2.6}$$

From Fig. 1 it is evident that

$$dS' = \frac{\mathbf{r}}{r} \cdot d\mathbf{S} = dS \cos\left(\frac{\mathbf{r}}{r}, d\mathbf{S}\right) \tag{2.7}$$

is the projection of the element of area $d\mathbf{S}$ on to the plane perpendicular to the radius vector at a given point. Consequently

$$\frac{dS'}{r^2} = d\Omega \tag{2.8}$$

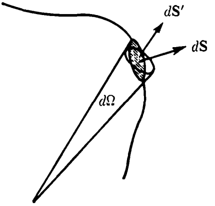


Fig. 1

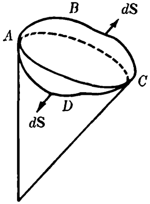


Fig. 2

is an element of the solid angle subtended by the surface element at the point where the charge e is situated. Using (2.7), (2.8), and (2.5), we can evaluate the integral (2.6)

$$N = \frac{e}{4\pi} \oint_S \frac{1}{r^2} \frac{\mathbf{r}}{r} \cdot d\mathbf{S} = \frac{e}{4\pi} \oint_S \frac{dS'}{r^2} = \frac{e}{4\pi} \int d\Omega = e \tag{2.9}$$

where we have used the fact that the total solid angle subtended by a closed surface at an internal point is 4π.

If the charge lies outside the closed surface S, the integral

$$\int_S \mathbf{D} \cdot d\mathbf{S}$$

will be positive when taken over ABC (Fig. 2) and negative (but equal in magnitude) when taken over ADC. Since the vector $d\mathbf{S}$ is always directed along the outward normal to the surface, the element of solid angle $d\Omega$ has a positive sign when it is subtended inside the surface and a negative sign when it is subtended outside the surface. It follows that the flux of the electric induction vector through a closed surface from a point charge outside the surface is equal to zero

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = 0 \tag{2.10}$$

If there are several point charges e_i, the electric induction vector $\mathbf{D}$ is equal to the sum of the induction vectors $\mathbf{D}_i$ due to the individual charges

$$\mathbf{D} = \sum_i \mathbf{D}_i \tag{2.11}$$

This fact is known from experiment and it is called the *principle of superposition.* Its principal mathematical consequence is the fact that the corresponding equations are linear. Hence, it follows from (2.11) that the equations satisfied by $\mathbf{D}$ must be linear, i.e., the sum of solutions is a solution, and a solution multiplied by a constant is also a solution.

Thus, for the flux of the vector $\mathbf{D}$ through an arbitrary closed surface surrounding point charges e_i, we obtain

$$N = \oint_S \mathbf{D} \cdot d\mathbf{S} = \sum_i \oint_S \mathbf{D}_i \cdot d\mathbf{S} = \sum_i e_i \tag{2.12}$$

In other words, the flux of electric induction through a closed surface is equal to the sum of the charges within the surface. This statement is known as *Gauss' electrostatic theorem.*

A continuous distribution of charge may be represented by an assembly of sufficiently small charges $\Delta q_i = \rho \Delta V_i$ inside volume elements ΔV_i into which a given volume may be divided. As $\Delta V_i \to 0$, these charges may be considered to be point charges. Gauss' theorem (2.12) can then be written in the form

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \lim_{\Delta V_i \to 0} \sum_i \rho \, \Delta V_i = \int_V \rho \, dV \tag{2.13}$$

where V is the volume enclosed by the closed surface S. It should be observed that (2.13) holds for an arbitrarily chosen region of space, bounded by a closed surface S.

Maxwell's Equation div $\mathbf{D} = \rho$. To obtain a differential relationship be-

tween the quantities which occur in (2.13), we must use Gauss' theorem (Appendix 2):

$$\oint_S \mathbf{A}\cdot d\mathbf{S} = \int_V \operatorname{div}\mathbf{A}\, dV \tag{2.14}$$

i.e., the flux of a vector through a closed surface is equal to the integral of the divergence of the vector over the volume enclosed by that surface. Applying this theorem to the left-hand side of equation (2.13), we obtain:

$$\oint_S \mathbf{D}\cdot d\mathbf{S} = \int_V \operatorname{div}\mathbf{D}\, dV \tag{2.15}$$

Hence, from (2.13), together with (2.15), it follows that

$$\int_V \{\operatorname{div}\mathbf{D} - \rho\}\, dV = 0 \tag{2.16}$$

This equation holds for an arbitrary volume of integration V however small. But if the integral of a function is equal to zero for an arbitrary region of integration, the function must vanish identically. This may easily be proved by considering the converse: if there exists a point at which the integrand is not equal to zero, then we can always choose, in the neighborhood of that point, some sufficiently small region within which, by virtue of its continuity, the function is not equal to zero and retains its sign. If we choose this small region as the region of integration, then, when we integrate, the integrand retains its sign and is not equal to zero. Consequently, the integral is not equal to zero in this region, which contradicts the hypothesis that it vanishes when taken over any region. Thus, the assumption that there exists a point at which the integrand is not equal to zero is false, and the converse is true.

Applying this theorem to equation (2.16), where the region of integration is arbitrary, we obtain

$$\operatorname{div}\mathbf{D} = \rho \tag{2.17}$$

This is one of Maxwell's equations and is readily recognized as the differential form of Gauss' theorem.

Electric Charges as the Sources and Sinks of D. The line of a vector **A**, defined in space as a *point function,* is a line such that its tangent at any point coincides with the direction of the vector at that point. We can thus define the lines of force of the vector **H**, current lines, lines of electric induction, and so on. From equation (2.14), it is clear that if $\operatorname{div}\mathbf{A} = 0$ in a volume V, then the flux of the vector **A** through a closed surface S enclosing V is equal to zero. Consequently, there are no sources or sinks of **A** in this volume, i.e., there are no points where the lines of **A** start or finish. But at those points where $\operatorname{div}\mathbf{A} \neq 0$, the lines of **A** either begin

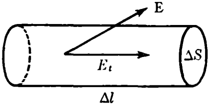
(div $\mathbf{A} > 0$) or end (div $\mathbf{A} < 0$), i.e., at points where div $\mathbf{A} \neq 0$ there are sources or sinks of $\mathbf{A}$.

Maxwell's equation (2.17) thus shows that electric charges are sources or sinks of the field of the electric induction $\mathbf{D}$: the lines of $\mathbf{D}$ begin on positive charges ($\rho > 0$) and end on negative charges ($\rho < 0$). Hence, we may say that the electric induction describes the creation of an electric field by charges.

§3. Ohm's Law and the Joule-Lenz Law in Differential Form

Differential Form of Ohm's Law. The potential difference between two points is numerically equal to the work done by the electric field when a unit positive charge is moved from one of the points to the other. Thus, if there exists a potential difference between two points in a conductor, then there exists an electric field between these points, producing the motion of electric charges.

Fig. 3



To obtain Ohm's law in differential form, we apply Ohm's law to an infinitely small cylinder in a conductor (Fig. 3). A current $\Delta I = j_t \Delta S$ flows along the axis of this cylinder, where ΔS is the area of the base of the cylinder, and j_t is the component of the current density along the axis. Since the cylinder is infinitely small, the electric field within it may be assumed constant and equal to $\mathbf{E}$. The potential difference $\Delta\varphi$ between the ends of the cylinder is

$$\Delta\varphi = E_t \Delta l$$

where E_t is the projection of the electric field intensity vector $\mathbf{E}$ on the axis of the cylinder. The potential difference is measured in volts and the field intensity in volts per meter. Then, for the infinitely small cylinder under consideration, Ohm's law may be written in the form

$$\Delta\varphi = \Delta I \Delta R = j_t \Delta S \Delta R \tag{3.1}$$

where ΔR is the (ohmic) resistance of the cylinder in ohms. The reciprocal

of the resistivity of the conductor is called its *conductivity* and is denoted by λ. The resistance ΔR of the cylinder is then

$$\Delta R = \frac{1}{\lambda}\frac{\Delta l}{\Delta S} \tag{3.2}$$

The conductivity has the dimensions

$$[\lambda] = \frac{1}{\text{ohm m}}$$

Substituting (3.2) in (3.1) and simplifying, we obtain

$$j_l = \lambda E_l \tag{3.3}$$

Since the direction of the axis of the cylinder was chosen arbitrarily, equation (3.3) holds for the projections of $\mathbf{j}$ and $\mathbf{E}$ in any arbitrary direction. Hence, we have the vector equation

$$\mathbf{j} = \lambda \mathbf{E} \tag{3.4}$$

which is called the differential form of Ohm's law. Equation (3.4) is one of the equations which supplement the electromagnetic field equations. There are no derivatives in this equation; nevertheless, it is called *differential*, since all the terms in it refer to the same point of the field.

Differential Form of the Joule-Lenz Law. The amount of heat Q, evolved in unit time in a conductor of resistance R, carrying a current I, is, according to the *Joule-Lenz law*, equal to

$$Q = I^2R \tag{3.5}$$

where Q is given in watts. Applying this law to an infinitely small cylinder (Fig. 3) the axis of which coincides with the direction of the current, we obtain

$$Q = (j\Delta S)^2 \frac{1}{\lambda}\frac{\Delta l}{\Delta S} \tag{3.6}$$

Since $\Delta S \cdot \Delta l = \Delta V$ is the volume of the cylinder, and $Q/dV = \tilde{q}$ is the quantity of the Joule heat evolved in unit volume in unit time, we find that

$$\frac{Q}{\Delta l \Delta S} = \tilde{q} = \frac{j^2}{\lambda} \tag{3.7}$$

$\tilde{q}$ is measured in watts per cubic meter. Remembering that $j^2 = \mathbf{j}\cdot\mathbf{j}$, and using equation (3.4), we may rewrite (3.7) in the form

$$\tilde{q} = \frac{j^2}{\lambda} = \mathbf{j}\cdot\mathbf{E} = \lambda E^2 \tag{3.8}$$

Any of these equations with $\tilde{q}$ in the left-hand-side represent the differential form of the Joule-Lenz law.

§4. Equation of Continuity and Displacement Current

Equation of Continuity. The *law of conservation of charge,* established experimentally, is expressed mathematically by the equation of continuity. The charge q contained within a volume V is given by the integral

$$q = \int_V \rho\, dV \tag{4.1}$$

If the charge q within a given volume is not constant, then there must be some motion of charge through the surface enclosing the volume. The amount of charge passing through this surface in time dt is clearly equal to

$$dt \int_S \mathbf{j} \cdot d\mathbf{S} \tag{4.2}$$

We take this value to be positive if the current flows out of the volume, and negative if the current flows into the volume. From the law of conservation of charge, it follows that such an inward or outward flow of charge must produce a corresponding change in the amount of charge contained in the volume. The change in time dt is equal to

$$dt \frac{dq}{dt} = dt \int_V \frac{\partial \rho}{\partial t}\, dV \tag{4.3}$$

Here, we use the fact that the derivative of an integral is equal to the integral of the derivative of the integrand, since the region of integration is independent of time. The sign of the derivative dq/dt is negative when the current flows out of the volume so that the value of (4.2) is positive. In the case of q increasing, the signs of (4.3) and (4.2) are reversed. It follows from the law of conservation of charge that the values of (4.3) and (4.2) are equal in absolute magnitude, but opposite in sign

$$\int_S \mathbf{j} \cdot d\mathbf{S} = -\int_V \frac{\partial \rho}{\partial t}\, dV \tag{4.4}$$

Applying Gauss' theorem (Appendix 1, Eq. A.2) to the left-hand side of this equation, we may rewrite it in the form

$$\int_V \left(\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} \right) dV = 0 \tag{4.5}$$

Since this equation holds for any arbitrary volume V, we conclude that the integrand is equal to zero

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} = 0 \tag{4.6}$$

This is the *equation of continuity,* which expresses the law of conservation of charge.

Conduction Current Lines. In the case of steady currents, the value of ρ at every point is constant, and, consequently

$$\frac{\partial \rho}{\partial t} = 0 \tag{4.7}$$

Hence, for constant currents, the equation of continuity takes the form

$$\operatorname{div} \mathbf{j} = 0 \tag{4.8}$$

This equation shows that constant current lines have neither beginning nor end. They are either closed lines, or else they go off to infinity. In the case of varying currents, the lines of the vector $\mathbf{j}$ are not closed, since, in this case, generally speaking

$$\operatorname{div} \mathbf{j} = -\frac{\partial \rho}{\partial t} \neq 0 \tag{4.9}$$

The lines of $\mathbf{j}$, therefore, start and finish at points where there is a change of charge density. The current density $\mathbf{j}$ is connected with the motion of charges, hence, it is called the *conduction current density* or the *transport current density*. Hence, we may say that the conduction current lines are not closed in the case of varying currents.

As an example, we may take an electric circuit containing a capacitor. It is well known that there cannot be a constant current round such a circuit, since charges cannot pass through a dielectric between the plates of the capacitor. Hence, the lines of the conduction current density $\mathbf{j}$ cannot be continued between the plates of the capacitor, and equation (4.8) is not satisfied.

Displacement Current. The situation is different in the case of the displacement current. Here, the presence of a capacitor in the circuit is no obstacle to the passage of a varying current. However, in this case too the charges cannot move through the dielectric between the capacitor plates. Hence, we must assume that some process takes place between the capacitor plates which is equivalent to a conduction current. We say that there is a displacement current between the plates, and that this completes the conduction current.

To obtain a mathematical expression for the displacement current, we differentiate both sides of Maxwell's equation

$$\rho = \operatorname{div} \mathbf{D}$$

with respect to time, obtaining

$$\frac{\partial \rho}{\partial t} = \operatorname{div} \frac{\partial \mathbf{D}}{\partial t} \tag{4.10}$$

Here we use the fact that the divergence is obtained by differentiation

with respect to the coordinates, and that the coordinates and time are independent variables, and therefore, the order of differentiation can be reversed. Substituting the expression for $\partial\rho/\partial t$ from (4.10) in the equation of continuity (4.6), we obtain

$$\operatorname{div}\left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}\right) = 0$$

It is clear from this that the lines of the vector

$$\mathbf{j}_{\text{total}} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}$$

are always closed. The vector $\mathbf{j}$ is the conduction current density, and the vector

$$\mathbf{j}_{\text{disp}} = \frac{\partial \mathbf{D}}{\partial t} \tag{4.11}$$

is called the *displacement current density.* The displacement current density has the same dimensions as the conduction current density, i.e.

$$\left[\frac{\partial \mathbf{D}}{\partial t}\right] = \text{amp/m}^2$$

Let us return to the example of an electric circuit with a capacitor. When a varying current I flows round such a circuit, the charge q on each plate of the capacitor varies: $I = dq/dt$. Let the area of the capacitor plate be S. The magnitude of the electric induction vector $\mathbf{D}$ between the plates is related to the charge q on the plate by the equation

$$D = \frac{q}{S}$$

Hence, it follows that

$$\frac{\partial D}{\partial t} = \frac{1}{S}\frac{\partial q}{\partial t} = \frac{I}{S}$$

is equal to the density I/S of the current which would flow between the capacitor plates if the space between them were entirely filled by a conductor.

In the physical sense, the displacement current

$$\mathbf{j}_{\text{disp}} = \frac{\partial \mathbf{D}}{\partial t} = \epsilon \frac{\partial \mathbf{E}}{\partial t} \tag{4.12}$$

has nothing in common with the conduction current. The displacement current density is a quantity proportional to the rate of change of the electric field at a given point, but this quantity is not called a *current* without good reason. According to (4.11), the displacement current is

accompanied by the same magnetic field which would appear in the presence of the corresponding conduction current. Thus, we may say that a change in the electric field produces a magnetic field, which is in accordance with the relationship between electric and magnetic fields, given by the law of electromagnetic induction. Not only is a change in the magnetic field always accompanied by an electric field, but, conversely, a change in the electric field is always accompanied by a magnetic field.

§5. Generalization of the Law of Total Current

The *law of total current* operates in the case of constant conduction currents. This law states that the circulation of the magnetic field intensity round a closed contour is equal to the algebraic sum of the currents in that contour. This law is expressed mathematically in the form

$$\oint_L \mathbf{H} \cdot d\mathbf{l} = I \tag{5.1}$$

where I is the algebraic sum of all the currents in an arbitrary closed contour L, i.e., the total conduction current in the contour L. The direction of integration round L and the direction of the total current make a right-hand screw system.

The law of total current can be obtained from the *Biot-Savart law* for an infinite straight-line current. The magnetic field of an infinite straight-line current, at an arbitrary distance r from it, is given by the formula

$$H = \frac{1}{2\pi R} \tag{5.2}$$

and the field $\mathbf{H}$ is directed along the tangent to a circle of radius r with its center on the axis of the current, in a plane perpendicular to the direction of the current. We shall consider the circulation of $\mathbf{H}$ around an arbitrary closed contour L in which there is a current, in a plane perpendicular to the direction of the current

$$\oint_L \mathbf{H} \cdot d\mathbf{l} \tag{5.3}$$

The integrand at some point of the contour has the form

$$\mathbf{H} \cdot d\mathbf{l} = H\, dl \cos(\mathbf{H}, d\mathbf{l}) \tag{5.4}$$

Remembering that, at every point r, $\mathbf{H}$ is directed along the tangent to the circle of radius r and center at the point of intersection of the direction of the current with the plane under discussion, we conclude that

$$dl \cos(\mathbf{H}, d\mathbf{l}) = dl_{\perp} \tag{5.5}$$

is the projection of the element dl along a direction perpendicular to the

radius vector $\mathbf{r}$. Hence, from the definition of the angle in radians, it follows that

$$\frac{dl_\perp}{r} = d\alpha \tag{5.6}$$

is the angle which dl subtends at the center of the circle. Using (5.2), we can then write down the equation

$$\mathbf{H}\cdot d\mathbf{l} = \frac{I}{2\pi r}\,dl_\perp = \frac{I}{2\pi}\,d\alpha \tag{5.7}$$

Now we may easily evaluate the integral round L

$$\oint_L \mathbf{H}\cdot d\mathbf{l} = \frac{I}{2\pi}\int d\alpha = I \tag{5.8}$$

If there is no current in L, then the integral (5.8) is clearly equal to zero. If there are several currents, then their magnetic field is equal to the sum of the fields produced by each current separately. Applying (5.8) to the sum of these fields, we obtain

$$\int_L \mathbf{H}\cdot d\mathbf{l} = \sum_i \oint_L \mathbf{H}_i\cdot d\mathbf{l}_i = \sum_i I_i = I \tag{5.9}$$

In (5.9) the sign of I_i depends on the direction of the current and the direction of integration round the contour L. If the direction of integration makes a right-hand screw system with the direction of the current I_i, then the sign of I_i is positive, otherwise it is negative. Therefore, in (5.9) I is the algebraic sum of the currents in the contour L, or, in other words, the total current in the contour. Thus, the law of total current is proved for infinite straight-line currents and for arbitrary contours in a plane perpendicular to the direction of the current. To free ourselves from these limitations, we write the law (5.9) in its differential form. The total current I in the contour L is clearly equal to

$$I = \int_S \mathbf{j}\cdot d\mathbf{S} \tag{5.10}$$

where S is the surface subtended by the contour L.

Equation (5.9) may be rewritten in the form

$$\oint_L \mathbf{H}\cdot d\mathbf{l} = \int_S \mathbf{j}\cdot d\mathbf{S} \tag{5.11}$$

Transforming the left-hand side in accordance with Stokes' theorem

$$\oint_L \mathbf{H}\cdot d\mathbf{l} = \int_S \operatorname{curl}\mathbf{H}\cdot d\mathbf{S} \tag{5.12}$$

(5.11) becomes

$$\int_S (\operatorname{curl} \mathbf{H} - \mathbf{j})\cdot d\mathbf{S} = 0 \tag{5.13}$$

Since the surface S is arbitrary, it follows that

$$\operatorname{curl} \mathbf{H} = \mathbf{j} \tag{5.14}$$

This is a differential relationship and is independent of how $\mathbf{j}$ behaves at other points. Therefore, although this relationship was deduced by considering straight-line currents, it is true for all currents. Let L be an arbitrary contour in which there is an arbitrary current I, and let S be a surface subtended by L. The current I flows through the surface S. We integrate both sides of equation (5.14) with respect to S

$$\int_S \operatorname{curl} \mathbf{H}\cdot d\mathbf{S} = \int_S \mathbf{j}\cdot d\mathbf{S} \tag{5.15}$$

Applying equations (5.12) and (5.10), we obtain expression (5.1) for an arbitrary current and contour. Thus, we have eliminated the restrictions under which we first obtained equation (5.11), and have proved the law of total current.

In the preceding paragraph it was pointed out that not only the conduction current, but also the displacement current produces a magnetic field, and that the field produced by the displacement current is equal to the field produced by the conduction current in accordance with equation (4.11). Hence, it is natural to generalize the law of total current (5.1), derived for the conduction current, by applying it to the displacement current. Consequently, in equation (5.1), I may be understood as the total current, equal to the sum of the conduction and displacement currents, and $\mathbf{j}$ in (5.15) must be replaced with the sum of the conduction and displacement current densities. The generalized form of these equations is, therefore

$$\int_L \mathbf{H}\cdot d\mathbf{l} = \int_S \left(\mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}\right)\cdot d\mathbf{S} \tag{5.16}$$

and hence

$$\operatorname{curl} \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \tag{5.17}$$

Equation (5.17) is one of Maxwell's differential equations.

From equation (5.1) it is evident that the magnetic field intensity is measured in amperes per meter

$$[H] = \text{amp/m}$$

The relationship of this quantity with the oersted, which is used to measure magnetic field intensity in the cgs emu system

$$1 \text{ amp/m} = 4\pi \times 10^{-3}\text{ Oe}$$

The proof of this relationship is given in the solution of Problem 16, Chapter I.

§6. Differential Form of the Law of Electromagnetic Induction

When the magnetic induction flux through a surface bounded by a closed conductor changes, an electric current is set up in the conductor by the induced emf, $\mathcal{E}^{\text{ind}}$. *Faraday's law of electromagnetic induction* is written in the form

$$\mathcal{E}^{\text{ind}} = -\frac{d\Phi}{dt} \tag{6.1}$$

where the minus sign takes into account the relationship between the direction of the induced emf and the rate of change of the flux.

The current appears in the conductor because an electric field is generated. The emf in a closed contour L is numerically equal to the work done in taking a unit positive charge round this contour, i.e.

$$\mathcal{E}^{\text{ind}} = \int_L \mathbf{E}\cdot d\mathbf{l} \tag{6.2}$$

The induced emf is measured in volts

$$[\mathcal{E}^{\text{ind}}] = V$$

By definition, the magnetic induction flux is given by

$$\Phi = \int_S \mathbf{B}\cdot d\mathbf{S} \tag{6.3}$$

The magnetic induction flux Φ is measured in webers

$$[\Phi] = \text{webers} = \text{volts/sec}$$

The weber is the magnetic flux through 1 m^2 of surface when the magnetic induction vector $\mathbf{B}$ is one tesla in magnitude and is perpendicular to the surface. One tesla equals 10^4 gauss

$$1 \text{ tesla} = 10^4 \text{ gauss}$$

The proof of this relationship is given in Problem 17, Chapter I. Taking (4.2) and (6.3) into account, equation (6.1) becomes

$$\oint_L \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\int_S \mathbf{B}\cdot d\mathbf{S} \tag{6.4}$$

The phenomenon of electromagnetic induction does not require the presence of a closed conductor. A change in the magnetic field is always

accompanied by the appearance of an electric field, irrespective of whether a conductor is present. A closed conductor merely allows the field to produce a current. Hence, (6.4) is true for any closed contour in space.

Applying Stokes' theorem to the left-hand side of equation (6.4) and remembering that the surface of integration on the right-hand side is independent of time (so that the derivative with respect to time may be included in the integrand) we obtain

$$\int_S \operatorname{curl} \mathbf{E} \cdot d\mathbf{S} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \tag{6.5}$$

Hence, since S is arbitrary, it follows that

$$\operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{6.6}$$

The minus sign shows that the vector which gives the rate of change of magnetic induction, $\partial \mathbf{B}/\partial t$, and the induced emf which it produces in a closed contour form a left-hand screw system, as shown in Fig. 4.

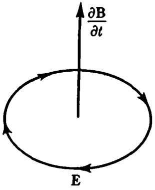


Fig. 4

Maxwell's equation (6.6) is the differential form of Faraday's law of electromagnetic induction.

§7. Maxwell's Equation, div B = 0

We apply the operator div to both sides of equation (6.6)

$$\operatorname{div} \operatorname{curl} \mathbf{E} = -\operatorname{div} \frac{\partial \mathbf{B}}{\partial t} \tag{7.1}$$

Since div curl $\equiv 0$, we obtain

$$0 = \operatorname{div} \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial}{\partial t} \operatorname{div} \mathbf{B}$$

Hence, div **B** is independent of time. Consequently, for any **B**, the value of div **B** is the same, including $\mathbf{B} = 0$. But for $\mathbf{B} = 0$, $\operatorname{div} \mathbf{B} = 0$. Therefore, $\operatorname{div} \mathbf{B} = 0$ for any **B**, i.e.,

$$\operatorname{div} \mathbf{B} = 0 \tag{7.2}$$

always.

This equation of Maxwell's is not independent, but is related to equation (6.6). It shows that the lines of the magnetic induction vector have neither beginning nor end. This means that there are no magnetic charges which could create a magnetic field in a manner similar to that by which electric charges create an electric field.

§8. Maxwell's System of Equations. The Energy of the Electromagnetic Field

Maxwell's System of Equations. Equations (2.17), (5.17), (6.6) and (7.2) make up Maxwell's system of equations

$$\left.\begin{aligned} \operatorname{curl} \mathbf{H} &= \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} && \text{(I)} \\ \operatorname{curl} \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && \text{(II)} \\ \operatorname{div} \mathbf{B} &= 0 && \text{(III)} \\ \operatorname{div} \mathbf{D} &= \rho && \text{(IV)} \end{aligned}\right\} \tag{8.1}$$

These equations must be supplemented by the constitutive relations

$$\mathbf{D} = \epsilon \mathbf{E} \qquad \mathbf{B} = \mu \mathbf{H} \qquad \mathbf{j} = \lambda \mathbf{E} \tag{8.1a}$$

Equations (8.1) and (8.1a) hold under the following conditions: (1) all matter in the field is stationary; (2) the values of ϵ, μ, λ for the medium are independent of time and of the field vectors; (3) there are no permanent magnets or ferromagnetic bodies in the field.

Completeness of the System. It has already been observed that equations (II) and (III) are not completely independent of each other. Owing to the mathematical identity div curl $\equiv 0$, equation (III) acts as an auxiliary condition in the solution of equation (II). Neither are equations (I) and (IV) completely independent. To prove this, we apply the operator div to equation (I)

$$\operatorname{div} \operatorname{curl} \mathbf{H} = \operatorname{div} \mathbf{j} + \frac{\partial}{\partial t} \operatorname{div} \mathbf{D} \tag{8.2}$$

Since div curl $\equiv 0$, we obtain

$$\frac{\partial}{\partial t} \operatorname{div} \mathbf{D} + \operatorname{div} \mathbf{j} = 0 \tag{8.3}$$

Comparing (8.3) with (4.6), we see that

$$\operatorname{div} \mathbf{D} = \rho \tag{8.4}$$

i.e., we obtain (IV). Thus, (I), (IV) and (8.1a) are independent, and using (8.1a), we can eliminate **D**, **B** and **j** from (I) and (II). Then we have two vector equations defining **E** and **H**. These two equations, together with the corresponding initial and boundary conditions, define **E** and **H** completely. These considerations, although not a strict proof, indicate that Maxwell's system of equations is a complete system.

The proof of the uniqueness of the solution of Maxwell's equations for a given distribution of charges and currents with given boundary and initial conditions is on the following lines. We assume that there are two distinct solutions. Since Maxwell's equations are linear, the difference of these two solutions is also a solution for zero currents and charges and zero initial and boundary conditions. Hence, using the expression for the energy of the electromagnetic field obtained at the end of this section, and the law of conservation of energy, we conclude that the difference of the solutions is identically equal to zero, i.e., that the solutions are equal and that for given conditions, the solution of Maxwell's equations is therefore unique.

Law of Conservation of Energy for the Electromagnetic Field. To be able to compare deductions from Maxwell's equations with experiment, we still need an expression for the energy of the electromagnetic field in terms of the field vectors. We therefore consider some volume V enclosed by a surface S. Within this volume, there is an electromagnetic field and currents, and Joule heat Q is evolved. By the law of conservation of energy (there is no doubt that this also applies to electromagnetic phenomena), the Joule heat is evolved at the expense of the energy of the field, since there are no other sources of energy available. Equation (3.8) leads to the relationship

$$Q = \int_V \mathbf{j} \cdot \mathbf{E}\, dV \tag{8.5}$$

Substituting the expression for **j** from equation (I), we find

$$Q = \int_V \mathbf{E} \cdot \operatorname{curl} \mathbf{H}\, dV - \int_V \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}\, dV \tag{8.6}$$

On the basis of formula (A.15) in Appendix 1, we have

$$\mathbf{E} \cdot \operatorname{curl} \mathbf{H} = \mathbf{H} \cdot \operatorname{curl} \mathbf{E} - \operatorname{div} \mathbf{E} \times \mathbf{H} \tag{8.7}$$

which, using equation (II) for curl **E**, allows equation (8.6) to be rewritten

$$Q = -\int_V \operatorname{div} \mathbf{E} \times \mathbf{H}\, dV - \int_V \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}\right) dV \tag{8.8}$$

We take into account the self-evident equalities

$$\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} = \frac{1}{2}\frac{\partial}{\partial t}(\mathbf{E}\cdot\mathbf{D})$$

$$\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{1}{2}\frac{\partial}{\partial t}(\mathbf{H}\cdot\mathbf{B})$$

and recall that V is independent of time. Then, taking out the operation of differentiation from under the integral sign in the second term of the right-hand side of equation (8.8), we obtain

$$\int_V \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}\right) dV = \frac{\partial}{\partial t}\frac{1}{2}\int_V (\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B})\, dV \equiv \frac{\partial W}{\partial t} \quad (8.9)$$

To transform the first integral of the right-hand side of equation (8.8), we use Gauss' theorem

$$\int_V \operatorname{div} \mathbf{E} \times \mathbf{H}\, dV = \int_S \mathbf{E} \times \mathbf{H}\, dS \equiv \int_S \mathbf{P}\cdot d\mathbf{S} \quad (8.10)$$

introducing the definition $\mathbf{P} = \mathbf{E} \times \mathbf{H}$.

Hence, equation (8.8) takes the form

$$\frac{\partial W}{\partial t} = -Q - \int_S \mathbf{P}\cdot d\mathbf{S} \quad (8.11)$$

This equation expresses the law of conservation of the energy for the electromagnetic field.

Expression for the Energy of the Electromagnetic Field and the Poynting Vector. The quantity

$$W = \frac{1}{2}\int_V (\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B})\, dV \quad (8.12)$$

is the energy of the electromagnetic field within the volume V. Equation (8.11) shows that the energy of the field in this volume changes due to two factors: the evolution of Joule heat Q (first term on the right-hand side of (8.11)) and the flow of energy through the surface S enclosing the volume (second term on the right-hand side of (8.11)). Since the second term on the right-hand side of (8.11) takes into account the flow of electromagnetic energy through the surface, it is clear that the vector

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \quad (8.13)$$

describes the motion of the electromagnetic energy in space. It is called the *Poynting vector.*

§9. Boundary Conditions

The quantities ϵ, μ, λ in Maxwell's equations (8.1), which describe the properties of the medium, are functions of the coordinates. ϵ, μ, and λ are not continuous functions everywhere; they are discontinuous on the boundaries separating different media. As is evident from equations (8.1) and (8.1a), the values of **E**, **H**, **D**, and **B** are discontinuous at points where ϵ, μ, and λ are discontinuous. Hence, we can say that the field vectors may be discontinuous on the boundaries separating different media. The conditions which describe the behavior of the field vectors on such boundaries are called the *boundary conditions*, and are deduced from Maxwell's equations. Here we must bear in mind the following important fact. Using Maxwell's equations to obtain the boundary conditions, we carry out a number of transformations by means of Gauss' and Stokes' theorems. But the applicability of these mathematical theorems depends on the continuity of the functions within the volume of integration, and in the cases under consideration the functions (the field vectors) within the volume of integration are discontinuous and the nature of these discontinuities still has to be determined. To avoid this difficulty, we proceed as follows. We suppose that instead of a boundary, where ϵ, μ and λ change discontinuously, there is a transition layer within which these values change very rapidly, but remain continuous. Hence, the field vectors also change very rapidly in the transition layer, but still remain continuous, and hence, the conditions of applicability of the mathematical theory are satisfied. After carrying out the necessary transformations, we let the thickness of the transition layer tend to zero, and thus, obtain the boundary conditions. To shorten the exposition, we shall not repeat the whole process of passing from the transition layer to the boundary surface every time.

The Normal Component of the Magnetic Induction $\mathbf{B}_n$. This condition is deduced using Maxwell's equation

$$\operatorname{div} \mathbf{B} = 0 \tag{9.1}$$

We consider a sufficiently small cylinder cut by the boundary surface separating two media, which we denote by the indices 1 and 2 (Fig. 5). We take the normal to the boundary surface directed toward the second medium. The bases of the cylinder have surfaces S_1 and S_2, parallel to the boundary surface. The area of the section of the cylinder cut by the boundary surface is denoted by S_0. Let the area of the lateral surface of the cylinder be S_{lat} and the height of the cylinder be h.

We integrate equation (9.1) over the volume of the cylinder

$$\int_V \operatorname{div} \mathbf{B} \, dV = 0 \tag{9.2}$$

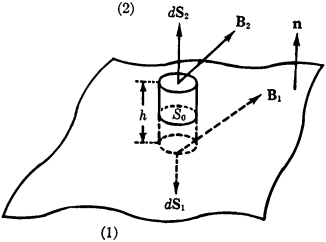


Fig. 5

Applying Gauss' theorem, we find

$$\int_V \operatorname{div} \mathbf{B}\, dV = \int_{S_1} \mathbf{B}\cdot d\mathbf{S} + \int_{S_2} \mathbf{B}\cdot d\mathbf{S} + \int_{S_{\text{lat}}} \mathbf{B}\cdot d\mathbf{S} = 0 \qquad (9.3)$$

The vector $d\mathbf{S}$ is directed along the normal $\mathbf{n}$ for integration over S_2, and in the opposite direction for integration over S_1. Since our cylinder is quite small, so that $\mathbf{B}$ may be considered constant for integration in one medium, we obtain

$$\begin{gathered} \int_{S_2} \mathbf{B}\cdot d\mathbf{S} = |\mathbf{B}_2| S_2 \cos(\mathbf{B}_2, \mathbf{n}) = B_{2n} S_2 \\ B_{2n} = |\mathbf{B}_2| \cos(\mathbf{B}_2, \mathbf{n}) \end{gathered} \qquad (9.4)$$

Here we take into account the fact that the direction of $d\mathbf{S}$ for the surface S_2 coincides with the chosen positive direction of the normal to the boundary surface. The indices of B_{2n} denote that we take the normal component of $\mathbf{B}$ in the second medium. In a similar manner, we calculate the integral over the surface S_1. However, for this surface, $d\mathbf{S}$ points in the negative direction

$$\int_{S_1} \mathbf{B}\cdot d\mathbf{S} = |\mathbf{B}_1| S_1 \cos(\mathbf{B}_1, -\mathbf{n}) = -B_{1n} S_1 \qquad (9.5)$$

The integral over the lateral surface is calculated using the *mean-value theorem*

$$\int_{S_{\text{lat}}} \mathbf{B}\cdot d\mathbf{S} = \langle B_{\text{lat}} \rangle S_{\text{lat}} \qquad (9.6)$$

where $\langle B_{\text{lat}} \rangle$ is the mean value of the induction vector on the lateral surface.

Using (9.4), (9.5) and (9.6), equation (9.3) may be rewritten

$$B_{2n} S_2 - B_{1n} S_1 + \langle B_{\text{lat}} \rangle S_{\text{lat}} = 0 \qquad (9.7)$$

Let the height of the cylinder h tend to zero. It is evident that, as $h \to 0$

$$S_2 \to S_0 \qquad S_1 \to S_0 \qquad S_{\text{lat}} \to 0 \tag{9.8}$$

Hence, in the limit, as $h \to 0$, we obtain

$$(B_{2n} - B_{1n})S_0 = 0 \tag{9.9}$$

and, since $S_0 \neq 0$,

$$B_{2n} = B_{1n} \tag{9.10}$$

i.e., the normal component of the magnetic induction is continuous on the boundary between the two media.

Using the fact that

$$B_{2n} = \mu_2 H_{2n} \qquad B_{1n} = \mu_1 H_{1n} \tag{9.11}$$

while μ_2, generally speaking, is not equal to μ_1, we see that the normal component of the magnetic field intensity is discontinuous on the boundary surface.

We further note that B_{2n} and B_{1n} are the normal components of $\mathbf{B}_2$ and $\mathbf{B}_1$, in the first and second media, respectively, along the same normal to the boundary surface.

The Normal Component of the Electric Induction D_n. This condition may be obtained using Maxwell's equation

$$\operatorname{div} \mathbf{D} = \rho \tag{9.12}$$

by an approach completely analogous to that of the preceding subsection, but using $\mathbf{D}$ instead of $\mathbf{B}$. After integrating (9.12) over the volume of the cylinder (Fig. 5), instead of (9.3), we obtain

$$\int_{S_2} \mathbf{D} \cdot d\mathbf{S} + \int_{S_1} \mathbf{D} \cdot d\mathbf{S} + \int_{S_{\text{lat}}} \mathbf{D} \cdot d\mathbf{S} = q \tag{9.13}$$

where q is the charge contained within the cylinder. All subsequent calculations are completely analogous to the previous case, and after proceeding to the limit $h \to 0$, instead of (9.9), we obtain the equation

$$(D_{2n} - D_{1n})S_0 = q \tag{9.14}$$

Here q denotes the surface charge on the boundary surface S_0. The quantity $\sigma = q/S_0$ is the surface charge density. Hence, the boundary condition for the normal components of the electric induction vector is

$$D_{2n} - D_{1n} = \sigma \tag{9.15}$$

and therefore, the normal component of $\mathbf{D}$ is discontinuous when there are surface charges on the surface boundary. These charges create an electric field, and therefore, $\mathbf{D}$ is discontinuous.

It should be noted that (9.15) also describes the behavior of the normal component of the electric field intensity

$$\epsilon_2 E_{n2} - \epsilon_1 E_{n1} = \sigma$$

The Tangential Component of the Electric Field Intensity. This condition is deduced using Maxwell's equation

$$\operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{9.16}$$

We suppose the boundary surface to be cut by a sufficiently small rectangular plane S, bounded by a contour L (Fig. 6). This plane cuts the boundary surface along the line l_0. Let the sides of the plane l_1 and l_2 be parallel to the boundary surface. We shall call the length of the sides of the plane which cut the surface l_{lat}. We integrate (9.16) over the surface S

$$\int_S \operatorname{curl} \mathbf{E} \cdot d\mathbf{S} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \tag{9.17}$$

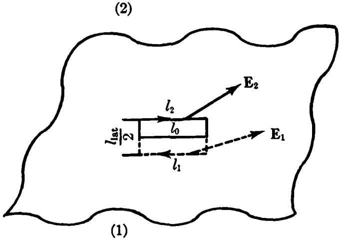


Fig. 6

The left-hand side of this equation may be transformed by Stokes' theorem

$$\int_S \operatorname{curl} \mathbf{E} \cdot d\mathbf{S} = \int_{l_2} \mathbf{E} \cdot d\mathbf{l} + \int_{l_1} \mathbf{E} \cdot d\mathbf{l} + \int_{l_{lat}} \mathbf{E} \cdot d\mathbf{l} \tag{9.18}$$

We take the positive direction of describing L to be that indicated in Fig. 6. Then we have

$$\int_{l_2} \mathbf{E} \cdot d\mathbf{l} = |\mathbf{E}_2| l_2 \cos(\mathbf{E}_2, d\mathbf{l}_2) = E_{2t} l_2 \tag{9.19}$$

Similarly

$$\int_{l_1} \mathbf{E} \cdot d\mathbf{l} = |\mathbf{E}_1| l_1 \cos(\mathbf{E}_1, -d\mathbf{l}_2) = -E_{1t} l_1 \tag{9.20}$$

The integral along l_{lat} is calculated with the aid of the mean-value theorem

$$\int_{l_{\mathrm{lat}}} \mathbf{E} \cdot d\mathbf{l} = \langle E_{\mathrm{lat}} \rangle l_{\mathrm{lat}} \tag{9.21}$$

We do not need to know the value of this integral more exactly, since, finally, the whole contour shrinks to the line l_0 and the value of the integral becomes zero.

To calculate the right-hand side of (9.17), we also use the mean-value theorem

$$\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} = \left\langle \frac{\partial B}{\partial t} \right\rangle S \tag{9.22}$$

Thus, using (9.18)–(9.22), equation (9.17) may be rewritten in the form

$$E_{2t} l_2 - E_{1t} l_1 + \langle E_{\mathrm{lat}} \rangle l_{\mathrm{lat}} = -\left\langle \frac{\partial B}{\partial t} \right\rangle S \tag{9.23}$$

Let l_{lat} tend to zero. Then, obviously, the whole surface S shrinks to the line l_0

$$l_2 \to l_0 \qquad l_1 \to l_0 \qquad l_{\mathrm{lat}} \to 0 \qquad S \to 0 \tag{9.24}$$

The values of $\langle E_{\mathrm{lat}} \rangle$ and $\langle \partial B/\partial t \rangle$ remain finite in this limiting process. Hence, in the limit, we obtain

$$(E_{2t} - E_{1t}) l_0 = 0$$

whence it follows that

$$E_{2t} = E_{1t} \tag{9.25}$$

Thus, the tangential component of the electric field intensity is continuous. However, the tangential component of the electric induction $\mathbf{D}$ is discontinuous.

The Tangential Component of the Magnetic Field Intensity. Starting from Maxwell's equation

$$\operatorname{curl} \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \tag{9.26}$$

and working in a manner similar to the previous case, we obtain in place of (9.17)

$$\int_S \operatorname{curl} \mathbf{H} \cdot d\mathbf{S} = \int_S \left(\mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot d\mathbf{S} \equiv I \tag{9.27}$$

and by repeating all the previous working, replacing $\mathbf{E}$ with $\mathbf{H}$, we obtain, in place of (9.23)

$$H_{2t} l_2 - H_{1t} l_1 + \langle H_{\mathrm{lat}} \rangle l_{\mathrm{lat}} = I \tag{9.28}$$

Let $l_{\mathrm{lat}} \to 0$, then

$$l_2 \to l_0 \qquad l_1 \to l_0 \qquad \langle H_{\mathrm{lat}} \rangle l_{\mathrm{lat}} \to 0$$

and the current I is expressed in terms of the current which flows along the surface and through l_0. Hence, in the limit, we obtain

$$(H_{2t} - H_{1t})l_0 = I_{\mathrm{surf}} \tag{9.29}$$

Remembering that $i_{\mathrm{surf}} = I_{\mathrm{surf}}/l_0$ is the surface current density, we may write the boundary condition for the tangential component of the magnetic field intensity

$$H_{2t} - H_{1t} = i_{\mathrm{surf}} \tag{9.30}$$

It must be stressed that i_{surf} is the surface current density in the direction perpendicular to that chosen for the tangential component of the magnetic field intensity.

If there is no surface current $i_{\mathrm{surf}} = 0$, the tangential component of the magnetic field intensity is continuous

$$H_{2t} = H_{1t} \tag{9.31}$$

The Tangential Component of the Current Density. This condition is obtained from the differential form of Ohm's law

$$\mathbf{j} = \lambda \mathbf{E} \tag{9.32}$$

Taking the projections of both sides of this equation in the tangential direction, we obtain

$$j_{2t} = \lambda_2 E_{2t} \qquad j_{1t} = \lambda_1 E_{1t}$$

Dividing the second equation by the first, term-by-term, and remembering that $E_{2t} = E_{1t}$, we find

$$\frac{j_{2t}}{j_{1t}} = \frac{\lambda_2}{\lambda_1} \tag{9.33}$$

Thus, if the electrical conductivities of the two media are different, there will be a different current density on each side of the surface of division.

The Normal Component of the Current Density. This condition is derived from the equation of continuity

$$\operatorname{div} \mathbf{j} = -\frac{\partial \rho}{\partial t}$$

Using an approach completely analogous to that in deducing the normal components of the magnetic and electric induction vectors, we obtain

$$j_{2n} - j_{1n} = -\frac{\partial \sigma}{\partial t}$$

where σ is the surface charge density. Consequently, the normal component of the current density is discontinuous only if the surface charge density changes with time.

PROBLEMS

1 By expressing curl and gradient in rectangular Cartesian coordinates, prove that

$$\operatorname{curl} \operatorname{grad} \varphi = 0$$

2 By expressing curl and divergence in rectangular Cartesian coordinates, prove that

$$\operatorname{div} \operatorname{curl} \mathbf{A} = 0$$

3 Calculate

$$\operatorname{grad} \varphi(r)$$

Solution:

$$\operatorname{grad} \varphi = \mathbf{i}\frac{\partial\varphi}{\partial x} + \mathbf{j}\frac{\partial\varphi}{\partial y} + \mathbf{k}\frac{\partial\varphi}{\partial z}$$

$$\frac{\partial\varphi}{\partial x} = \frac{\partial\varphi}{\partial r}\frac{\partial r}{\partial x} = \varphi'\frac{\partial r}{\partial x}$$

and

$$r = \sqrt{x^2 + y^2 + z^2}$$

Hence

$$\frac{\partial r}{\partial x} = \frac{2x}{2\sqrt{x^2 + y^2 + z^2}} = \frac{x}{r}$$

The derivatives $\partial r/\partial y$ and $\partial r/\partial z$ are calculated in a similar manner. Consequently

$$\operatorname{grad} \varphi(r) = \frac{d\varphi}{dr}(\mathbf{i}x + \mathbf{j}y + \mathbf{k}z) = \frac{d\varphi}{dr}\frac{\mathbf{r}}{r}$$

In particular, if $\varphi(r) = r$, then

$$\operatorname{grad} r = \frac{\mathbf{r}}{r}$$

If $\varphi(r) = 1/r$, then

$$\operatorname{grad}\frac{1}{r} = -\frac{1}{r^2}\frac{\mathbf{r}}{r}$$

4 Prove that grad φ is perpendicular to the surface $\varphi = \text{const}$.

Proof: We have

$$d\varphi = \frac{\partial\varphi}{\partial x}dx + \frac{\partial\varphi}{\partial y}dy + \frac{\partial\varphi}{\partial z}dz = \operatorname{grad}\varphi\cdot d\mathbf{r}$$

where $d\mathbf{r} = \mathbf{i}\,dx + \mathbf{i}\,dy + \mathbf{k}\,dz$. Let $d\mathbf{r}$ lie along the surface $\varphi = \text{const}$, then, obviously, $d\varphi = 0$ and

$$\operatorname{grad}\varphi\cdot d\mathbf{r} = 0$$

This means that grad φ is perpendicular to $d\mathbf{r}$. Q.E.D.

5 Calculate the divergence of the radius vector $\mathbf{r}$.
Solution:

$$\operatorname{div} \mathbf{r} = \frac{\partial \mathbf{r}_x}{\partial x} + \frac{\partial \mathbf{r}_y}{\partial y} + \frac{\partial \mathbf{r}_z}{\partial z} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$$

6 Calculate $\operatorname{div} r \cdot \mathbf{A}$, where $\mathbf{A}$ is a constant vector.

Answer: $\dfrac{\mathbf{r} \cdot \mathbf{A}}{r}$

7 Calculate $\operatorname{div} \boldsymbol{\omega} \times \mathbf{r}$, where $\boldsymbol{\omega}$ is a constant vector.

Answer: 0

8 Calculate $\operatorname{div} \left(\dfrac{\mathbf{r}}{r}\right)$.

Answer: $\dfrac{2}{r}$

9 Find $\operatorname{div} \mathbf{A} \times (\mathbf{r} \times \mathbf{B})$, where $\mathbf{A}$ and $\mathbf{B}$ are constant vectors.

Answer: $2\mathbf{A} \cdot \mathbf{B}$

10 Calculate $\operatorname{curl} \varphi(r)\mathbf{r}$.

Answer: 0

11 Calculate $\operatorname{curl} \boldsymbol{\omega} \times \mathbf{r}$, where $\boldsymbol{\omega}$ is a constant vector.

Answer: $2\boldsymbol{\omega}$

12 Calculate the circulation of the vector $\boldsymbol{\omega} \times \mathbf{r}$ around a circle L of radius r_0 lying in a plane perpendicular to the constant vector $\boldsymbol{\omega}$. The center of the circle lies at the origin of coordinates.
Solution: $\boldsymbol{\omega} \times \mathbf{r}_0$ lies along the tangent to the circle at every point. Consequently

$$\oint_L \boldsymbol{\omega} \times \mathbf{r} \cdot d\mathbf{l} = \omega r_0 \int_L dl = 2\pi\omega r_0^2$$

We assume that the direction of describing the circle is such that the vectors $(\boldsymbol{\omega} \times \mathbf{r})$ and $d\mathbf{l}$ coincide. If we change the direction, we change the sign of the integral.

This problem may also be solved using Stokes' theorem. We have

$$\oint_L \boldsymbol{\omega} \times \mathbf{r} \cdot d\mathbf{l} = \int_S \operatorname{curl} \boldsymbol{\omega} \times \mathbf{r} \cdot d\mathbf{S}$$

where S is the area of the circle.

Using the result of Problem 11, we obtain

$$\int_L \boldsymbol{\omega} \times \mathbf{r} \cdot d\mathbf{l} = 2\omega \int_S dS = 2\omega\pi r_0^2$$

13 Calculate the flux of the radius vector through the surface shown in Fig. 7.
Solution: We take the origin of coordinates at the center of the base of the cylinder and the z axis along the axis of the cylinder. We have

$$\int_S \mathbf{r} \cdot d\mathbf{S} = \int_{S_l} \mathbf{r} \cdot d\mathbf{S} + \int_{S_u} \mathbf{r} \cdot d\mathbf{S} + \int_{S_{\text{lat}}} \mathbf{r} \cdot d\mathbf{S}$$

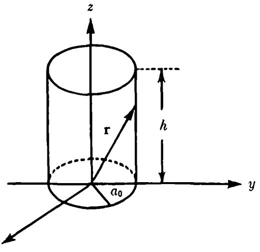


Fig. 7

where S_l, S_u and S_{lat} are, respectively, the surfaces of the lower and upper ends, and of the lateral surface. We have

$$\int_{S_l} \mathbf{r}\cdot d\mathbf{S} = 0$$

since $\cos(\mathbf{r}, d\mathbf{S}) = 0$

$$\int_{S_u} \mathbf{r}\cdot d\mathbf{S} = h\pi a_0^2$$

since $r\cos(\mathbf{r}, d\mathbf{S}) = h$

$$\int_{S_{lat}} \mathbf{r}\cdot d\mathbf{S} = a_0 2\pi a_0 h$$

since $r\cos(\mathbf{r}, d\mathbf{S}) = a_0$.

Finally

$$\int_S \mathbf{r}\cdot d\mathbf{S} = 3\pi a_0^2 h$$

This problem may also be solved using Gauss' theorem. Using the result of Problem 12, we obtain

$$\int_S \mathbf{r}\cdot d\mathbf{S} = \int_V \operatorname{div}\mathbf{r}\, dV = 3\pi a_0^2 h$$

14 Determine the dimensions of certain quantities using Maxwell's equations, Coulomb's law, and Faraday's law of electromagnetic induction.

For example, if we know the dimensions of the volume density

$$[\rho] = \text{coul}/m^3$$

and using Maxwell's equation

$$\operatorname{div}\mathbf{D} = \rho$$

we obtain the following expression for the dimensions of the electric induction

$$[D] = \text{coul}/m^2$$

15 Using the law of total current in the absolute Gaussian system

$$\int \mathbf{H}\cdot d\mathbf{l} = \frac{4\pi}{c} I \tag{1}$$

and the fact that 1 coul = 3×10^9 cgs esu express the unit of magnetic field intensity in the SI system (1 amp/m) in oersteds.

Solution: Using formula (1), we calculate the field at a distance r from an infinite straight line conductor along which a current I flows. We have

$$H = \frac{4\pi}{c}\frac{I}{2\pi r} \tag{2}$$

Here, H is in oersteds, I is in cgs esu, r is in centimeters and c in cm/sec. We have

$$I \text{ cgs} = 3 \times 10^9 I \text{ amp}$$

where I(amp) is the current in amperes. Hence, (2) may be written in the form

$$H(\text{Oe}) = \frac{4\pi}{3 \times 10^{10}} \frac{3 \times 10^9}{10^2} \frac{I(\text{amp})}{2\pi r(m)} \tag{3}$$

where $r(m)$ is the distance in meters.

On the other hand, from the law of total current (5.1), written in SI units, we obtain, instead of (2)

$$H \text{ amp/m} = \frac{I \text{ amp}}{2\pi r \text{ m}} \tag{4}$$

From (3) and (4) it follows that

$$H(\text{Oe}) = 4\pi\, 10^{-3} H \text{ amp/m}$$

This means that

$$1 \text{ amp/m} = 4\pi\, 10^{-3} \text{ Oe}$$

16 Using the result of Problem 15, find the relationship between the SI unit of magnetic induction, the tesla, and the unit of magnetic induction in the absolute Gaussian system, the gauss.

Solution: In the SI system, we have

$$B(\text{T}) = \mu' \mu_0 H \text{ amp/m} \tag{5}$$

where μ_0 is given by (1.4). From Problem 15, it follows that

$$H\text{amp/m} = \frac{1}{4\pi} 10^3 \text{ Oe}$$

By the definition of induction in the Gaussian system, we may write

$$B(\text{G}) = \mu' H(\text{Oe})$$

Hence, instead of (5), we may write

$$B(\text{T}) = \mu_0 \frac{1}{4\pi} 10^3 B(\text{G}) = 10^{-4} B(\text{G})$$

from which it follows that

$$1 \text{ T} = 10^4 \text{ G}$$

CHAPTER 2

Electrostatics

§10. Possibility of Considering Electrical and Magnetic Problems Separately

Electrostatics is the study of the constant electric field of stationary charges. Mathematically, the subject of electrostatics is distinguished from the whole subject of electromagnetic theory by the following requirements: (1) all quantities are constant with respect to time; (2) there is no motion of charges, i.e., $\mathbf{j} = 0$.

Under these conditions, Maxwell's equations and the boundary conditions take the following form

$$\begin{aligned} \operatorname{curl} \mathbf{H} &= 0 & \qquad \operatorname{curl} \mathbf{E} &= 0 \\ \operatorname{div} \mathbf{B} &= 0 & \qquad \operatorname{div} \mathbf{D} &= \rho \\ B_{2n} - B_{1n} &= 0 & \qquad D_{2n} - D_{1n} &= \sigma \\ H_{2t} - H_{1t} &= 0 & \qquad E_{2t} - E_{1t} &= 0 \end{aligned}$$

Thus, the equations split into two groups of independent equations, one of which contains terms relating to the magnetic field only, and the other, terms relating to the electric field only. The electrostatic and magnetostatic fields may, therefore, be discussed quite independently. This is possible only when the fields are independent of time. In the general case of fields which vary with respect to time, the electric and magnetic fields cannot be separated in this way, but must be discussed together.

§11. Electrostatic Field in a Homogeneous Medium

We shall consider the electrostatic field in a homogeneous medium (ϵ = const). In empty space, $\epsilon = \epsilon_0$.

Fundamental Problems of Electrostatics. The equations of the electrostatic field and the boundary conditions are

$$\begin{aligned} \operatorname{curl} \mathbf{E} &= 0 \qquad & E_{2n} - E_{1n} &= \frac{\sigma}{\epsilon} \\ \epsilon \operatorname{div} \mathbf{E} &= \rho \qquad & E_{2t} - E_{1t} &= 0 \end{aligned} \tag{11.1}$$

There are three fundamental problems in the theory of the electrostatic field: (1) for a given electric field, i.e., for a known value of the intensity **E** as a function of the coordinates $\mathbf{E} = \mathbf{E}(x, y, z)$, to find the charge distribution, i.e., to find ρ and σ as functions of the coordinates; (2) for a given charge distribution, i.e., with given functions for ρ and σ, to find the value of **E** at all points of space; (3) to find the forces acting on charges in the electrostatic field.

The solution of the first of these problems is trivial. Rewriting (11.1) in the form

$$\rho = \epsilon \operatorname{div} \mathbf{E} \qquad \sigma = \epsilon(E_{2n} - E_{1n}) \tag{11.2}$$

gives the solution immediately. The second and third problems are of greatest interest in electrostatics.

Potential Nature of the Electrostatic Field. A *potential vector field* is one in which the curl is identically equal to zero. The electrostatic field is a potential field, since

$$\operatorname{curl} \mathbf{E} = 0 \tag{11.3}$$

In the electrostatic field, the work done by the field when a charge is displaced from one point to another is independent of the path taken and depends on only the initial and final points of this path. This follows directly from (11.3). Let us take two different paths Γ and Γ' connecting

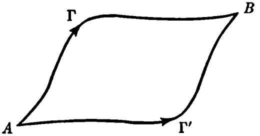


Fig. 8

the points A and B (Fig. 8). We shall consider the work done by the field when a unit positive charge moves round the closed contour formed by Γ and Γ' (Fig. 8). This is equal to

$$\oint_{\Gamma, -\Gamma'} \mathbf{E} \cdot d\mathbf{l} = \int_S \operatorname{curl} \mathbf{E} \cdot d\mathbf{S} = 0 \tag{11.4}$$

where S is a surface subtended by the contour. In (11.4) we use Stokes' theorem and (11.3). Thus

$$\int_{\Gamma, -\Gamma'} \mathbf{E}\cdot d\mathbf{l} = \int_{\Gamma} \mathbf{E}\cdot d\mathbf{l} + \int_{-\Gamma'} \mathbf{E}\cdot d\mathbf{l} = \int_{\Gamma} \mathbf{E}\cdot d\mathbf{l} - \int_{\Gamma'} \mathbf{E}\cdot d\mathbf{l} = 0 \quad (11.5)$$

i.e.

$$\int_{\Gamma} \mathbf{E}\cdot d\mathbf{l} = \int_{\Gamma'} \mathbf{E}\cdot d\mathbf{l} \quad (11.6)$$

Q.E.D. The paths Γ and Γ' are completely arbitrary.

Scalar Potential. The fact that when a charge moves from one point to another the work done by the field is independent of the path taken, means that there exists a scalar function φ such that the difference between its values at the initial and final points of the path defines the work done. φ is called the *scalar potential.*

Since curl grad $\equiv 0$, the general solution of (11.3) is

$$\mathbf{E} = -\operatorname{grad} \varphi \quad (11.7)$$

where the minus sign is of historical origin, but has no basic significance. Due to the presence of the minus sign in (11.7), the electric field vector points in the direction of decreasing potential. On the basis of (11.7), the expression for the integral (11.6) may be written in the form

$$\int_A^B \mathbf{E}\cdot d\mathbf{l} = \int_A^B - \operatorname{grad} \varphi\cdot d\mathbf{l} = -\int_A^B d\varphi = \varphi(A) - \varphi(B) \quad (11.8)$$

In (11.8) we take into account the fact that

$$d\varphi = \frac{\partial\varphi}{\partial x}\,dx + \frac{\partial\varphi}{\partial y}\,dy + \frac{\partial\varphi}{\partial z}\,dz = \operatorname{grad} \varphi\cdot d\mathbf{l} \quad (11.9)$$

since the components of the displacement $d\mathbf{l}$ are dx, dy, and dz. Equation (11.8) shows that the work done when a charge is displaced from one point to another is, in fact, expressed by the potential difference between these two points.

Normalization of the Potential. The potential itself is an auxiliary quantity. Its numerical value has no physical meaning, and cannot be measured experimentally. The only thing which has a physical meaning is the potential difference, which can be measured experimentally. This difference remains unchanged if we add the same constant to the value of the potential at all points in space, since in measuring the potential difference this arbitrary constant is eliminated. This constant is completely arbitrary, and we may choose it at our own discretion. Thus, we may make the potential at any fixed point equal to a predetermined value, and then the potential at all other points is defined in terms of this single value. The procedure of making the scalar potential single-valued is called *normaliza-*

tion of the potential. In practical electrical work, the usual method of normalizing is to take the potential of the earth to be zero. In theoretical physics, we usually take the potential at infinity to be zero, provided all charges lie in a finite region of space.

Normalizing the potential $\varphi(\infty) = 0$, and taking B at infinity in equation (11.8), we obtain

$$\varphi(A) = \int_A^\infty \mathbf{E} \cdot d\mathbf{l} \tag{11.10}$$

where the path of integration is arbitrary and is chosen according to the requirements of the particular problem.

Potential of a Point Charge. The field of a point charge e is spherically symmetrical and therefore, the potential of a point charge also possesses spherical symmetry. It depends only on the distance r between the point at which the potential is measured and the point charge generating that potential. Recalling that the intensity $\mathbf{E}$ of the field of e at a distance r from it is given by

$$\mathbf{E} = \frac{1}{4\pi\epsilon} \frac{e}{r^2} \frac{\mathbf{r}}{r} \tag{11.11}$$

and using (11.10), we obtain

$$\varphi(r) = \int_r^\infty \mathbf{E} \cdot d\mathbf{l} = \frac{1}{4\pi\epsilon} \frac{e}{r} \tag{11.12}$$

In calculating (11.12), we choose the path of integration along the radius vector. Thus, the potential of a point charge e at a distance r is directly proportional to e/r. Conversely, if we know φ as a function of the coordinates, it is not difficult to determine the electric field intensity using (11.7).

Potential of a System of Point Charges. If we have two point charges e_1 and e_2, each of which in isolation produces fields $\mathbf{E}_1$ and $\mathbf{E}_2$, respectively, the total field $\mathbf{E}$ produced by the two charges together, is equal to

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = -\operatorname{grad} \varphi_1 - \operatorname{grad} \varphi_2 = -\operatorname{grad} \varphi$$

where $\varphi = \varphi_1 + \varphi_2$. Thus, the potential of a system of point charges is equal to the sum of the potentials produced by each of the charges in isolation. Hence, the potential of a system of charges e_i equals

$$\varphi = \frac{1}{4\pi\epsilon} \sum_i \frac{e_i}{r_i}$$

where r_i is the distance from the charge e_i to the point at which the potential is measured. If the coordinates of the point are (x, y, z), and the coordinates of the charge e_i are (x_i, y_i, z_i), then this formula may be written in the more detailed form

$$\varphi(x, y, z) = \frac{1}{4\pi\epsilon} \sum_i \frac{e_i}{\sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}} \tag{11.13}$$

Potential of a Continuous Distribution of Charges. In the case of a continuous charge distribution of density ρ, the whole volume may be divided up into infinitesimal elements ΔV_i containing charge $\rho_i \Delta V_i$. In the limit, as $\Delta V_i \to 0$, we may apply equation (11.13) for the potential of a system of point charges. As a result, we obtain

$$\begin{aligned}\varphi(x, y, z) &= \lim_{\Delta V_i \to 0} \frac{1}{4\pi\epsilon} \sum_i \frac{\rho_i \Delta V_i}{r_i} \\ &= \frac{1}{4\pi\epsilon} \int_V \frac{\rho(x', y', z')\, dx'\, dy'\, dz'}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}} \qquad (11.14)\end{aligned}$$

If the surface charge density on a surface S is σ, then, by analogy to the previous case, we have

$$\varphi = \frac{1}{4\pi\epsilon} \int_S \frac{\sigma\, dS}{r} \tag{11.15}$$

Finite Nature of the Potential in the Case of a Continuous Distribution of Charge with Finite Density. Equation (11.14) for a continuous distribution of charge is obtained by generalizing equation (11.13) for a system of point charges; however, there is an essential difference between these formulas. Equation (11.13) for a point charge gives an infinite value of the potential at the point where the charge is situated ($1/r \to \infty$ as $r \to 0$). The potential (11.14) for a continuous distribution of charge with constant density has a finite value at all points. This becomes clear if we use (11.14) to calculate the potential at the point (x, y, z), taking it as the origin of coordinates (i.e., $x = 0, y = 0, z = 0$). The calculation may similarly be carried out in a spherical system of coordinates, in which an element of volume is given by

$$dx'\, dy'\, dz' = r'^2 \sin\theta'\, d\theta'\, d\alpha'\, dr'$$

Then (11.14) takes the form

$$\varphi(0, 0, 0) = \frac{1}{4\pi\epsilon} \int_V \rho(r', \alpha', \theta') r' \sin\theta'\, d\theta'\, d\alpha'\, dr'$$

Hence, it is evident that if ρ is finite at all points (as was assumed), and the charges are all in a finite region of space, then the potential φ is, in fact, finite at all points of space.

The infinite nature of the potential φ at the site of a point charge is due to the infinite value of the charge density, since we consider the volume

containing the charge to be equal to zero. In a finite distribution of charge, however, in an infinitely small volume, there is an infinitely small amount of charge, since the charge density is finite at all points. Hence, the potential φ is finite at all points of space.

If we have both surface and volume charges, then (11.14) and (11.15) may be written as the single equation

$$\varphi = \frac{1}{4\pi\epsilon}\int_V \frac{\rho\, dV}{r} + \frac{1}{4\pi\epsilon}\int_S \frac{\sigma\, dS}{r} \tag{11.16}$$

However, this direct method of calculating the potential is not always convenient, since it sometimes leads to very complicated calculations. Furthermore, the applicability of this formula must be subjected to special analysis if the charges are not all in a finite region of space. In these cases, it is better to reduce the problem of finding the potential to the solution of a differential equation.

§12. Laplace's Equation and Poisson's Equation

Deduction of the Equations. To obtain the differential equation satisfied by φ, we substitute in the equation

$$\operatorname{div} \mathbf{E} = \frac{\rho}{\epsilon} \tag{12.1}$$

the expression for the electric field intensity in terms of the potential

$$\mathbf{E} = -\operatorname{grad} \varphi \tag{12.2}$$

We shall use the formula

$$\operatorname{div} \operatorname{grad} \varphi = \nabla^2\varphi \tag{12.3}$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{12.4}$$

is the Laplace operator.

Hence, (12.1) becomes

$$\nabla^2\varphi = -\frac{\rho}{\epsilon} \tag{12.5}$$

This is *Poisson's equation.* In regions of space where $\rho = 0$, this reduces to *Laplace's equation*

$$\nabla^2\varphi = 0 \tag{12.6}$$

Properties of the Potential. In the preceding section, it was shown that the potential φ is finite everywhere. It is also clear that the derivatives

of φ with respect to the coordinates must also be finite everywhere, since, if they were infinite, it would imply that the electric field was infinite, and this has no physical meaning.† However, since the derivatives are finite, φ must be continuous. Thus, we can say that the potential φ is a continuous finite function with finite derivatives with respect to the coordinates. This is the necessary condition for the differential equation (12.5) to have a solution.

In the preceding section it was shown that equation (11.6) is the general expression for the potential created by volume and surface charges in a finite region of space. It is clear that the expression for φ must satisfy equation (12.5), as may be shown by direct substitution of (11.6) in (12.5).

As we have said, in many cases it is more convenient to find the potential by solving Poisson's equation (12.5). This is due to the fact that when the potential is introduced using (12.2), then (11.3) is immediately satisfied, and the problem of finding three functions $E_x(x, y, z)$, $E_y(x, y, z)$, and $E_z(x, y, z)$ is reduced to that of finding a single function $\varphi(x, y, z)$. Another important advantage of the use of Poisson's equation to find the potential is its wider range of application. Equation (11.6) assumes that all charges lie in a finite region of space, and hence, the normalization of the potential at infinity to zero has meaning. But Poisson's equation does not assume the normalization of the potential and the absence of charges at infinity.

Infinite Charged Plate. As an example of the use of Laplace's equation, let us consider the potential due to an infinite plate of thickness a, charged with constant volume density $\rho \equiv$ const. It is assumed that the permittivity at all points of space, including the plate, is constant and equal to ϵ. We choose our system of coordinates in such a way that the x and y axes lie in the central plane of the plate. It is clear that the potential depends only on the z coordinate, since in the x and y directions, the plate is infinite and uniformly charged. Hence, Poisson's equation (12.5) assumes the following form

$$\begin{gathered}
\frac{d^2\varphi_1}{dz^2} = 0 \quad \text{at} \quad z \leqslant -\frac{a}{2} \\
\frac{d^2\varphi_2}{dz^2} = -\frac{\rho}{\epsilon} \quad \text{at} \quad -\frac{a}{2} < z < \frac{a}{2} \\
\frac{d^2\varphi_3}{dz^2} = 0 \quad \text{at} \quad z > \frac{a}{2}
\end{gathered} \tag{12.7}$$

† But **E** may be discontinuous, in which case the derivatives of φ are discontinuous. Further, φ itself may be discontinuous: for example, if we have a surface layer of electric dipoles, a so-called *double layer*.

The solutions are written in the form

$$\varphi_1 = A_1z + B_1$$

$$\varphi_2 = -\frac{1}{2}\frac{\rho}{\epsilon}z^2 + A_2z + B_2 \tag{12.8}$$

$$\varphi_3 = A_3z + B_3$$

As the normalization condition, we take the potential in the central plane to be equal to zero $\varphi_z(0) = 0$. From (12.8) it then follows that

$$B_2 = 0 \tag{12.9}$$

is a necessary condition.

Further, we remember that, due to the symmetry, the field in the central plane $z = 0$ is equal to zero. This means that

$$\left.\frac{\partial\varphi_2}{\partial z}\right|_{z=0} = 0, \quad \text{i.e.,} \quad A_2 = 0 \tag{12.10}$$

We now use the condition that the potential and its derivative are continuous at $z = \pm a/2$, i.e.

$$\begin{aligned} \varphi_1\left(-\frac{a}{2}\right) &= \varphi_2\left(-\frac{a}{2}\right) \qquad & \left.\frac{\partial\varphi_1}{\partial z}\right|_{z=-\frac{a}{2}} &= \left.\frac{\partial\varphi_2}{\partial z}\right|_{z=-\frac{a}{2}} \\ \varphi_2\left(\frac{a}{2}\right) &= \varphi_3\left(\frac{a}{2}\right) \qquad & \left.\frac{\partial\varphi_2}{\partial z}\right|_{z=+\frac{a}{2}} &= \left.\frac{\partial\varphi_3}{\partial z}\right|_{z=+\frac{a}{2}} \end{aligned} \tag{12.11}$$

This gives

$$A_1 = \frac{\rho}{\epsilon}\frac{a}{2} \qquad B_1 = \frac{1}{2}\frac{\rho}{\epsilon}\frac{a^2}{4}$$

$$A_3 = -\frac{\rho}{\epsilon}\frac{a}{2} \qquad B_3 = \frac{1}{2}\frac{\rho}{\epsilon}\frac{a^2}{4}$$

Therefore, the required solution of (12.8) becomes

$$\varphi_1 = \frac{\rho}{\epsilon}\frac{a}{2}\left(z + \frac{a}{4}\right) \quad \text{at} \quad z < -\frac{a}{2}$$

$$\varphi_2 = -\frac{1}{2}\frac{\rho}{\epsilon}z^2 \quad \text{at} \quad -\frac{a}{2} < z < \frac{a}{2} \tag{12.12}$$

$$\varphi_3 = \frac{\rho}{\epsilon}\frac{a}{2}\left(-z + \frac{a}{4}\right) \quad \text{at} \quad z > \frac{a}{2}$$

and hence, for the field $E_z = -\partial\varphi/\partial z$, we obtain the following expressions

$$E_z = -\frac{\partial \varphi_1}{\partial z} = -\frac{\rho}{\epsilon}\frac{a}{2} \quad \text{at} \quad z < -\frac{a}{2}$$

$$E_z = -\frac{\partial \varphi_2}{\partial z} = \frac{\rho}{\epsilon} z \quad \text{at} \quad -\frac{a}{2} < z < \frac{a}{2} \tag{12.13}$$

$$E_z = -\frac{\partial \varphi_3}{\partial z} = \frac{\rho}{\epsilon}\frac{a}{2} \quad \text{at} \quad z > \frac{a}{2}$$

Thus, from the central plane to the surface of the plate the field increases linearly with distance, and acts in opposite directions on the opposite sides of the central plane. On the surfaces of the plate, the field reaches its maximum absolute value. Outside the plate, the field is constant and equal to this maximum value.

Infinite Uniformly Charged Circular Cylinder. As a further example of the use of Poisson's equation, we shall consider the potential of an infinite cylinder of radius a, uniformly charged with volume density ρ = const.

We take the z axis along the axis of the cylinder. Since the charge distribution possesses axial symmetry, so does the potential φ, i.e., $\varphi = \varphi(r)$. Hence, we use a cylindrical system of coordinates, where the axial angle is α. In cylindrical coordinates, Laplace's equation becomes

$$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial r^2} + \frac{1}{r}\frac{\partial \varphi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\varphi}{\partial \alpha^2} + \frac{\partial^2\varphi}{\partial z^2} \tag{12.14}$$

Since the potential φ depends on r only, in the present case (12.14) becomes

$$\nabla^2\varphi = \frac{d^2\varphi}{dr^2} + \frac{1}{r}\frac{d\varphi}{dr} = \frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi}{dr}\right) \tag{12.15}$$

Hence, Poisson's equation (12.15) may be rewritten

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi_1}{dr}\right) = -\frac{\rho}{\epsilon} \quad \text{at} \quad 0 < r < a$$

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\varphi_2}{dr}\right) = 0 \quad \text{at} \quad r > a \tag{12.16}$$

The general solution of this equation is found by direct integration

$$\varphi_1 = -\frac{1}{4}\frac{\rho}{\epsilon}r^2 + A_1 \ln r + B_1 \tag{12.17}$$

$$\varphi_2 = A_2 \ln r + B_2$$

Since the potential must be finite at all points, and $\log r \to -\infty$ as $r \to 0$, we must put $A_1 = 0$ in (12.17). A suitable normalization of the potential is $\varphi_1(0) = 0$. Under this condition

$$B_1 = 0 \tag{12.18}$$

The condition that the potential and its derivative are continuous at $r = a$ gives two algebraic equations for the unknowns A_2 and B_2

$$\left.\begin{aligned} A_2 \ln a + B_2 &= -\frac{1}{4}\frac{\rho}{\epsilon}a^2 \\ \frac{A_2}{a} &= -\frac{1}{2}\frac{\rho}{\epsilon}a \end{aligned}\right\} \tag{12.19}$$

Hence, it follows that

$$\left.\begin{aligned} \varphi_1(r) &= -\frac{1}{4}\frac{\rho}{\epsilon}r^2 \qquad \text{at} \quad 0 < r < a \\ \varphi_2(r) &= \frac{1}{2}\frac{\rho}{\epsilon}a^2 \ln\frac{a}{r} - \frac{1}{4}\frac{\rho}{\epsilon}a^2 \qquad \text{at} \quad r > a \end{aligned}\right\} \tag{12.20}$$

The electric field intensity is expressed by the formulas

$$\left.\begin{aligned} E_r &= -\frac{\partial\varphi_1}{\partial r} = \frac{1}{2}\frac{\rho}{\epsilon}r \qquad \text{at} \quad 0 < r < a \\ E_r &= -\frac{\partial\varphi_2}{\partial r} = \frac{1}{2}\frac{\rho}{\epsilon}\frac{a^2}{r} \qquad \text{at} \quad r > a \end{aligned}\right\} \tag{12.21}$$

Thus, inside the cylinder the potential increases in direct proportion to the radius, while outside the cylinder it decreases in inverse proportion to the radius.

§13. Conductors in an Electrostatic Field

Absence of an Electrostatic Field Inside a Conductor. Bodies in which the presence of an electric field causes a movement of charge, i.e., an electric current, are called conductors. Mathematically, conductors are characterized by the fact that their conductivity $\lambda \neq 0$. Since in electrostatics we consider the case of stationary charges, when $\mathbf{j} = 0$, from equation (8.1a)

$$\mathbf{j} = \lambda\mathbf{E} = 0 \tag{13.1}$$

it follows immediately that when there is no movement of charge in the conductor, i.e., in the case of electrostatic equilibrium

$$\mathbf{E} = 0 \tag{13.2}$$

It must be stressed that the field in the conductor is equal to zero only when there are no currents in it, and the charges are in equilibrium. When there are currents in the conductor, the electric field, causing the current flow, is different from zero.

Absence of Volume Charges in a Conductor. Since, in the case of electrostatic equilibrium, there is no field inside a conductor, then

$$\operatorname{div} \mathbf{D} = 0 \tag{13.3}$$

It follows, therefore, from the equation

$$\operatorname{div} \mathbf{D} = \rho \tag{13.4}$$

that the volume charge density inside a conductor is equal to zero

$$\rho = 0 \tag{13.5}$$

The charges in a conductor are concentrated on the surface in a layer of atomic thickness. Physically, the situation is as follows. If a conductor is charged, then due to the repulsive forces of like charges, the latter distribute themselves over the surface in such a way that the field within the conductor is equal to zero. If the conductor is then placed in an external electrostatic field, the charges on the surface redistribute themselves so that the field inside the conductor, which is the sum of the external field and the field created by the charges on the surface of the conductor, is once again equal to zero. This redistribution of the charges on the surface of a conductor placed in an external electrostatic field is called *electrostatic induction.*

Field Close to the Surface of a Conductor. The electric field close to the surface of a conductor may be found from the boundary conditions. Taking the outward normal to the surface of the conductor to be positive, the boundary conditions are

$$\begin{aligned} \epsilon E_{2n} - \epsilon_0 E_{1n} &= \sigma \\ E_{2t} &= E_{1t} \end{aligned} \tag{13.6}$$

where the index 2 denotes the space outside the conductor, and 1, the space inside the conductor. In equation (13.6), we assume that the permittivity of a conductor is approximately equal to the permittivity of empty space. Since the electric field is zero inside a conductor

$$E_{1n} = E_{1t} = 0 \tag{13.7}$$

we obtain from (13.6) the expression for the components of field outside the conductor

$$E_{2n} = \frac{\sigma}{\epsilon} \qquad E_{2t} = 0 \tag{13.8}$$

Thus, the field outside the conductor near its surface is directed along the outward normal and is equal to σ/ϵ in absolute value

$$\mathbf{E} = \frac{\sigma}{\epsilon}\mathbf{n} \tag{13.9}$$

The absence of a tangential component of the field close to the surface of a conductor is evident: such a component must lead to the motion of charges along the surface of the conductor. Equilibrium is reached when there is no such motion, i.e., when the tangential component is equal to zero.

Potential of a Conductor. From a condition wherein the field $\mathbf{E}$ inside a conductor is zero, it follows that the potential is constant over the whole conductor, since if $\varphi(A)$ and $\varphi(B)$ are the potentials of the points A and B on the conductor, then, on the basis of (11.8)

$$\varphi(A) - \varphi(B) = \int_A^B \mathbf{E}\cdot d\mathbf{l} = 0, \quad \text{i.e.,} \quad \varphi(A) = \varphi(B) \tag{13.10}$$

Hence, one may speak of the *potential of a conductor.*

The potential of a conductor depends on the form of the conductor and the amount of charge on it, and also on the charge distributions on other conductors in the neighborhood.

Capacitance of a Conductor. If we take an isolated conductor, i.e., a conductor at a sufficiently great distance from other bodies and charges so that there is nothing to cause a redistribution of the charges on the conductor, then its potential depends on only its shape and the charge. The ratio of the charge q on an isolated conductor to its potential φ is called its *capacitance* C

$$C = \frac{q}{\varphi} \tag{13.11}$$

If the conductor is not isolated, then its potential depends on the shape, the charges and the distribution of other conductors. The capacitance of a conductor is measured in farads. From (13.11) it follows that the dimensions of a capacitance are

$$[C] = F = \text{coul/V}$$

In the cgs esu system, the capacitance is measured in centimeters, and the formula for the capacitance is identical with (13.11). Since 1 volt = $\frac{1}{300}$ cgs esu, 1 coul = 3×10^9 cgs esu, it follows directly from (13.11) that

$$1F = 9 \times 10^{11}\ \text{cm}$$

Metal Screen. The electric field at points inside a conductor is the sum of the fields: (1) the field produced by the surface charges of the conductor, and (2) the external electric field in which the conductor is situated. The surface charges on the conductor always rearrange themselves in such a way as to compensate the external electric field, so that the field inside the conductor becomes zero. As we have observed, the surface charges are situated on the surface of the conductor in a layer of atomic thickness.

We shall assume the permittivity of the conductor to be equal to the permittivity of empty space. If we cut away the whole interior of the conductor, except for the thin boundary layer in which the surface charges are concentrated, then nothing has been changed in this picture: the surface charges on the conducting boundary layer shell are distributed in such a way that the field in the cavity is equal to zero. Such a conducting shell is called a *screen*. It screens off the inside of the cavity from the fields outside; the field inside the cavity is independent of the fields outside the shell.

We now consider a cavity in an infinite conducting medium. If charges are placed inside this cavity, charges of opposite sign will be induced on the surface of the cavity. It will later be proved that the total value of the charge induced on the surface of the cavity is equal in absolute magnitude to the charge inside the cavity. A charge of the same value, and of the same sign, flows away through the conducting medium to infinity. In the electrostatic case, the field inside the conducting medium must be equal to zero. This field is caused by the charges inside the cavity and the charges induced on the surface of the cavity. If we now remove the conducting medium, leaving a metal shell surrounding the cavity, with the induced surface charges on the shell, then the overall picture does not change at all, and the field outside the shell is, as before, equal to zero. Thus, a conducting shell, connected by a conductor to an infinitely distant point ("an earthed shell"), screens the space outside the shell from the charges inside the cavity. The field outside the shell is independent of the fields inside the cavity.

It has been stated previously that the charge induced on the inside surface of a screen is equal and opposite to the charge contained in the cavity. To prove this statement, we apply Gauss' theorem to the volume contained within the cavity, taking the inside surface of the screen as the surface of integration (Fig. 9)

$$\int_S \mathbf{E}\cdot d\mathbf{S} = \frac{1}{\epsilon_0}\int_V \rho\, dV \tag{13.12}$$

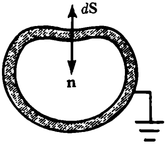


Fig. 9

On the surface of a conductor

$$\mathbf{E} = \frac{\sigma}{\epsilon_0}\mathbf{n} \tag{13.13}$$

where $\mathbf{n}$ is the normal to the surface of the conductor pointing inward into the cavity. Consequently

$$\mathbf{E}\cdot d\mathbf{S} = \frac{\sigma}{\epsilon_0}\mathbf{n}\cdot d\mathbf{S} = -\frac{\sigma}{\epsilon_0}\,dS \qquad (\cos(\mathbf{n}, d\mathbf{S}) = -1) \tag{13.14}$$

Hence, (13.12) assumes the form

$$-\int_S \sigma\,dS = \int_V \rho\,dV \tag{13.15}$$

Q.E.D.

Experiments show that a metal screen does not have to be continuous. Screens made of fine-meshed metal grids give sufficiently good screening.

Screening shells are used very widely in practice. The lead shielding shell of a power cable acts as a screen so that the field of the conductors in the cable is confined to the cable, and hence, the electrostatic effect of the cable cannot affect neighboring communication transmission lines.

Capacitors. A *capacitor* is a system of two conductors such that they completely screen the space between them. The conductors are called the *capacitor plates*. The plates of a capacitor carry equal and opposite charges. Between the plates there is a definite difference in potential, proportional to the charge on the plates. The ratio of the charge on each plate q to the potential difference between the plates $\varphi_1 - \varphi_2$ is called the capacitance C

$$C = \frac{q}{\varphi_1 - \varphi_2} \tag{13.16}$$

The units of capacitance have already been discussed in connection with (13.11).

The simplest capacitor is a parallel-plate capacitor consisting of two parallel conducting plates separated by a distance that is much less than the linear dimensions of the plates. The statement that the distance between the plates is small is necessary for the condition of screening of the space between the plates to be satisfied as completely as possible. The greater the ratio of the linear dimensions of the plates to the distance between them, the more complete will be the screening.

Cylindrical and spherical conductors are also widely used. A cylindrical capacitor consists of two coaxial conducting cylinders such that the distance between their walls is much less than their height. A spherical capacitor consists of two concentric conducting spheres. In a spherical capacitor,

the condition of the total screening of the internal field is satisfied completely.

As an example of a calculation of the capacitance, we shall consider a spherical capacitor with an internal sphere of radius R_1, and an external sphere of radius R_2. The space between the plates is filled with a dielectric of permittivity ϵ. Let the charge on the capacitor (i.e., the magnitude of the charge on each plate) be q. We imagine a sphere of radius R $(R_1 \leqslant R \leqslant R_2)$ and apply Gauss' theorem

$$\int_S \mathbf{D}\cdot d\mathbf{S} = q \tag{13.17}$$

Since the field is spherically symmetrical, and the vector $\mathbf{D}$ lies along the radius vector drawn from the center of the sphere, we obtain

$$\int_S \mathbf{D}\cdot d\mathbf{S} = D \int_S dS = D\,4\pi R^2 \tag{13.18}$$

Hence, using (13.17), it follows that

$$E = \frac{1}{4\pi\epsilon}\frac{q}{R^2} \tag{13.19}$$

This means that a uniformly charged sphere creates in the outside space a field that would be produced if all the charge on the sphere were concentrated at the center. In calculating the potential difference, it is convenient to take the path of integration along the radius vector drawn from the center of the sphere. Then we obtain

$$\varphi_1 - \varphi_2 = \int_{(1)}^{(2)} \mathbf{E}\cdot d\mathbf{l} = \int_{R_1}^{R_2} E\,dR = \frac{q}{4\pi\epsilon}\int_{R_1}^{R_2}\frac{dR}{R^2} = \frac{q}{4\pi\epsilon}\frac{R_2 - R_1}{R_1R_2} \tag{13.20}$$

Calculating the capacitance according to (13.16), we obtain

$$C = 4\pi\epsilon\frac{R_1R_2}{R_2 - R_1} \tag{13.21}$$

Similar methods may be used to calculate the capacitance in other cases. This is recommended as an exercise (see problems).

System of Conductors. If we have several conductors, then the potential of each conductor is not related to the charge on that conductor by the simple formula (13.11). In this case, the potential of a conductor depends on the charge, the shape and the positions of all the other conductors. If we have N conductors with potentials φ_i and charges q_i, then, since the potential of the sum of charges is equal to the sum of the potentials of the individual charges, we may write down the following expression for the potentials of the individual conductors

$$\varphi_i = \sum_{j=1}^{N} \alpha_{ij}q_j \qquad i = 1, 2, \ldots, N \tag{13.22}$$

The coefficients α_{ij} are called the *coefficients of potential.* They depend on the shapes and dimensions of the conductors, on their positions with respect to each other, and on the permittivity of the medium. The theoretical evaluation of these coefficients is generally a difficult mathematical problem. Often these coefficients are determined experimentally.

The coefficients of potential α_{ij} are not independent of each other. We may show this as follows. Let σ_j be the surface charge density on the j^{th} conductor. Then we can write down the following expression for the potential of the i^{th} conductor

$$\varphi_i = \frac{1}{4\pi\epsilon_0}\sum_{j=1}^{n}\int_{S_j}\frac{\sigma_j\,dS_j}{r_{ij}} \tag{13.23}$$

where r_{ij} is the distance between the element of the surface of integration of the j^{th} conductor and some arbitrary fixed point of the i^{th} conductor. The sum (13.23) includes also the term for $i = j$.

The charge on the i^{th} conductor is equal to

$$q_i = \int_{S_i}\sigma_i\,dS_i \tag{13.24}$$

Let the charges on the conductors change, so that the charge on the i^{th} conductor becomes equal to

$$q'_i = \int_{S_i}\sigma'_i\,dS_i \tag{13.25}$$

Multiplying both sides of (13.23) by q'_i and summing over all values of i

$$\sum_i q'_i\varphi_i = \frac{1}{4\pi\epsilon_0}\sum_{i,j}\int_{S_i}\int_{S_j}\frac{\sigma'_i\sigma_j\,dS_i\,dS_j}{r_{ij}}$$

$$= \frac{1}{4\pi\epsilon_0}\sum_{i,j}\int_{S_j}\sigma_j\,dS_j\int\frac{\sigma'_i\,dS_i}{r_{ij}} = \sum_j q_j\varphi'_j \tag{13.26}$$

where the order of integration is reversed because the integration is performed with respect to different independent variables. Thus, we obtain the following relationship

$$\sum_i q'_i\varphi_i = \sum q_i\varphi'_i \tag{13.27}$$

which is called the *reciprocity theorem.* From the reciprocity theorem we can immediately obtain the condition satisfied by the coefficients of potential α_{ij}. Assume that at first the charges on all the conductors except the k^{th} are equal to zero. Then (13.22) assumes the form

$$\varphi_i = \alpha_{ik}q_k \tag{13.28}$$

Next, assume that the charges on all the conductors except the l^{th} are equal to zero. In this case, (13.22) becomes

$$\varphi'_i = \alpha_{il} q'_l \tag{13.29}$$

The reciprocity theorem (13.27) assumes the form

$$q'_l \alpha_{lk} q_k = q_k \alpha_{kl} q'_l \tag{13.30}$$

Hence

$$\alpha_{lk} = \alpha_{kl} \tag{13.31}$$

This is the required condition for the coefficients of potential a_{ij}.

The system of equations (13.22) may be solved for the charges q_i

$$q_i = \sum_{j=1}^{n} C_{ij} \varphi_j \tag{13.32}$$

where $C_{ij} = A_{ij}/D$. Here, D is the determinant of the coefficients of the system of equations (13.22), and A_{ij} is the cofactor of the element a_{ij} in this determinant. From (13.31) we may conclude that the coefficients C_{ij} satisfy the following condition

$$C_{ij} = C_{ji} \tag{13.33}$$

The coefficients C_{ij} for $i \neq j$ are called *coefficients of induction*, and for $i = j$, *coefficients of capacitance* or simply capacitances. Since a positive charge on an insulated conductor produces a positive field, the coefficients of capacitance are always positive. We shall now consider two conductors, one of which is earthed, and the other of which insulated and positively charged. This positive charge induces a negative charge on the earthed conductor by electrostatic induction. Hence, using (13.32), we may conclude that the coefficients of induction are negative or, in some cases, equal to zero, and that they cannot be positive.

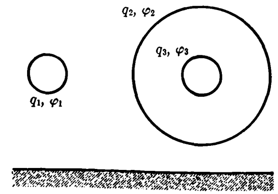


Fig. 10

As an example, we shall consider three conducting spheres (Fig. 10). The charges and potentials of these three conductors are denoted, respec-

tively, by q_1, φ_1; q_2, φ_2 and q_3, φ_3. To determine the values of C_{ij}, we take equations (13.32), which, in the present case, may be written

$$\left.\begin{aligned} q_1 &= C_{11}\varphi_1 + C_{12}\varphi_2 + C_{13}\varphi_3 \\ q_2 &= C_{21}\varphi_1 + C_{22}\varphi_2 + C_{23}\varphi_3 \\ q_3 &= C_{31}\varphi_1 + C_{32}\varphi_2 + C_{33}\varphi_3 \end{aligned}\right\} \tag{13.34}$$

To determine the coefficients, we must take a number of situations and determine the corresponding values of q_i and φ_i. From the equations obtained, we can then evaluate the coefficients C_{ij}.

We shall assume that $q_3 = 0$ and that the second sphere is earthed. Clearly, in this case, $\varphi_3 = \varphi_2 = 0$, and consequently, (13.34) assumes the form

$$q_1 = C_{11}\varphi_1 \qquad q_2 = C_{21}\varphi_1 \qquad 0 = C_{31}\varphi_1 \tag{13.35}$$

Hence, it follows that $C_{31} = C_{13} = 0$, i.e., the coefficient of induction between two conductors screened from each other is equal to zero.

We shall now assume that the first and second spheres are earthed, i.e., $\varphi_1 = 0$, $\varphi_2 = 0$, $q_3 \neq 0$. Then (13.34) takes the form

$$q_1 = 0 \qquad q_2 = C_{23}\varphi_3 \qquad q_3 = C_{33}\varphi_3 \tag{13.36}$$

In the discussion of a metal screen, it has been shown that a charge induced on the inside surface of an earthed conducting shell is equal and opposite to the charge within the screen. Consequently, $q_2 = -q_3$. From (13.36), we have

$$C_{23} = -C_{33} \tag{13.37}$$

Thus, the coefficient of induction between two conductors when one of them completely encloses the other is equal and opposite in sign to the coefficient of capacitance of the inner conductor. A system of two conductors, one of which completely screens the other, is called a capacitor. The capacitance of the capacitor is the coefficient of capacitance of the inner conductor, which is equal and opposite to the coefficient of induction between the conductors. We have already discussed the calculation of the capacitance of capacitors.

Method of Images. One very important method of solving a number of problems of electrostatics is the method of images which consists, in essence, of the following. The problem is to find the potential of an electric field. If the potential is known, then the field is found by taking the gradient of the potential. The distribution of the potential in space is described by the shape of equipotential surfaces, i.e., the surfaces on which the potential has a constant value. The electric field intensity is directed along the normal to the equipotential surface at any given point.

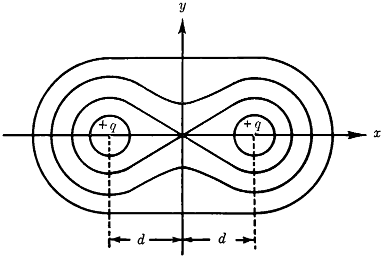


Fig. 11

Finding the shape of the equipotential surfaces of a system of point charges is very easy in principle. We shall consider, for example, two positive point charges q a distance $2d$ apart (Fig. 11). Since the potential of a point charge at a distance r from it is given by $\varphi = q/4\pi\epsilon_0 r$, the potential, at the point (x, y, z), of the two point charges is defined by

$$\varphi(x, y, z) = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{\sqrt{(x-d)^2+y^2+z^2}} + \frac{1}{\sqrt{(x+d)^2+y^2+z^2}}\right) \quad (13.38)$$

From (13.38) we obtain the equation for the equipotential surfaces

$$\frac{1}{\sqrt{(x-d)^2+y^2+z^2}} + \frac{1}{\sqrt{(x+d)^2+y^2+z^2}} = \text{const} \quad (13.39)$$

Each equipotential surface is characterized by a corresponding value of the potential $\varphi_1, \varphi_2 \ldots$

Fig. 11 shows the lines of intersection of the xy plane with the equipotential surfaces. The same equipotential surfaces are obtained by rotating the picture given in Fig. 11 about the x axis.

Let a conducting insulated surface coincide with an equipotential surface of potential φ_0. If this conductor carries a charge $2q$, then its potential is defined as φ_0. The potential at all points outside the conductor is given by (13.38). Thus, finding the field produced by a charged conductor is reduced to finding the field produced by two equal like point charges.

The equation of the equipotential surfaces of two unlike point charges is given by a formula similar to (13.38), but with a minus sign in the second term. The form of the equipotential surfaces in this case is shown in Fig. 12. The potential along the y axis is equal to zero, and, therefore, it is equal to zero in the plane $x = 0$.

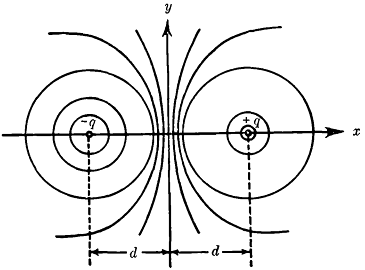


Fig. 12

Thus, if we replace a point charge $-q$ by an infinite plane conducting surface $x = 0$ carrying a charge $-q$, there is no change in the picture of the equipotential surfaces in the half-space $x > 0$, and, therefore, there is no change in the electric field. Thus, the field in the half-space $x > 0$ in the presence of a point charge $+q$ and an infinite conducting plane $x = 0$, is the same as the field caused by a point charge $+q$ and another point charge $-q$ at a point which is a mirror image with respect to $x = 0$ of the site of the first charge. There is no difficulty in determining the field of two point charges. This method of solution is called the *method of images.* Basically, the method consists of choosing a distribution of charges such that one of the equipotential surfaces coincides with the surface of the conductor under consideration.

We shall now determine the field of a charge $+q$ situated at a point $x = d$ in the presence of a conducting plane $x = 0$. In accordance with the previous discussion, the potential at all points $x > 0$ is given by the formula

$$\varphi = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{\sqrt{(x-d)^2 + y^2 + z^2}} - \frac{1}{\sqrt{(x+d)^2 + y^2 + z^2}}\right) \tag{13.40}$$

The electric field in the plane $z = 0$ is equal to

$$\begin{aligned} E_x &= -\frac{\partial\varphi}{\partial x} = \frac{q}{4\pi\epsilon_0}\left\{\frac{x-d}{[(x-d)^2 + y^2]^{3/2}} - \frac{x+d}{[(x+d)^2 + y^2]^{3/2}}\right\} \\ E_y &= -\frac{\partial\varphi}{\partial y} = \frac{q}{4\pi\epsilon_0}\left\{\frac{y}{[(x-d)^2 + y^2]^{3/2}} - \frac{y}{[(x+d)^2 + y^2]^{3/2}}\right\} \end{aligned} \tag{13.41}$$

In the plane $x = 0$, the component E_y vanishes, and the component E_x is equal to

$$E_x = \frac{-q}{2\pi\epsilon_0} \frac{d}{(y^2 + d^2)^{3/2}} \tag{13.42}$$

In accordance with the boundary condition for D_n, the surface charge density on the surface $x = 0$ of the conductor is given by the expression

$$\sigma = \frac{-q}{2\pi} \frac{d}{(y^2 + d^2)^{3/2}} \tag{13.43}$$

The force between the point charge q and the conducting surface $x = 0$ is equal to the force between the charge and its image, i.e.

$$F = \frac{-q^2}{16\pi\epsilon_0 d^2} \tag{13.44}$$

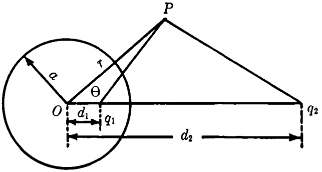


Fig. 13

We shall now consider the equipotential surfaces of two charges which differ in absolute value. For convenience we shall introduce a system of polar coordinates with their center at the point O (Fig. 13). The positions of the charges q_1 and q_2 are given by the coordinates $\theta_1 = 0$, $r_1 = d_1$; $\theta_2 = 0$, $r_2 = d_2$, respectively. The potential at P is equal to

$$\varphi(r, \theta) = \frac{1}{4\pi\epsilon_0}\left(\frac{q_1}{\sqrt{r^2 + d_1^2 - 2rd_1\cos\theta}} + \frac{q_2}{\sqrt{r^2 + d_2^2 - 2rd_2\cos\theta}}\right) \tag{13.45}$$

It is easy to verify that when $d_1 = a^2/d_2$ ($a < d_2$) and $q_1 = -(a/d_2)q_2$, the equality $\varphi(a, \theta) = 0$ holds, i.e., under these conditions, the potential on a sphere of radius a is equal to zero. Consequently, this equipotential surface may be replaced by an earthed sphere. Thus, if we have an earthed conducting sphere of radius a and a point charge q_2 outside it, at a distance d_2 from the center of the sphere, the field outside the sphere is the same as the field set up by the charge q_2 and its image, the charge $q_1 = -(a/d_2)q_2$, placed at the point a^2/d_2, $\theta = 0$ inside the sphere. The force between the charge and the sphere is

$$F = \frac{q_1 q_2}{4\pi\epsilon_0(d_2 - d_1)^2} = -\frac{d_2 a q_2^2}{4\pi\epsilon_0(d_2^2 - a^2)^2} \tag{13.46}$$

§14. Dielectrics in an Electrostatic Field

Dipole Moment. A pair of equal and opposite point charges is called a dipole (Fig. 14). The electrical properties of a dipole are described in the first approximation by the dipole moment. If $\mathbf{l}$ is a vector pointing from the negative to the positive charge of the dipole, equal in magnitude to the distance between the charges, and q is the absolute magnitude of each of the charges of the dipole, then the vector

$$\mathbf{p} = q\mathbf{l} \tag{14.1}$$

is called the *dipole moment.*

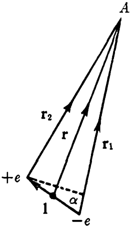


Fig. 14

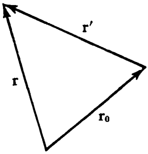


Fig. 15

We shall consider the electric field of a dipole at distances many times greater than the distance between the charges. At such distances it is governed, in the first approximation, by the dipole moment. The potential φ due to the dipole at some point A (Fig. 14) is equal to the sum of the potentials due to the charges of the dipole

$$\varphi = \frac{q}{4\pi\epsilon}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) = \frac{q}{4\pi\epsilon}\,\frac{r_1 - r_2}{r_1 r_2} \tag{14.2}$$

Taking the origin of coordinates halfway between the charges (Fig. 14), we have

$$\mathbf{r}_1 = \mathbf{r} + \frac{\mathbf{l}}{2} \qquad \mathbf{r}_2 = \mathbf{r} - \frac{\mathbf{l}}{2}$$

Squaring this equation, and ignoring terms of the order of l^2 in comparison with r_1^2 and r_2^2, we find

$$r_1^2 \approx r^2 + \mathbf{r}\cdot\mathbf{l}$$

$$r_2^2 \approx r^2 - \mathbf{r}\cdot\mathbf{l}$$

Hence, it is clear that the equation

$$r_1^2 - r_2^2 = (r_1 + r_2)\cdot(r_1 - r_2) = 2\mathbf{r}\cdot\mathbf{l}$$

is valid to terms of the order of l^2, and consequently, to the same degree of accuracy, we have

$$r_1 - r_2 = \frac{2}{r_1 + r_2}\mathbf{r}\cdot\mathbf{l}$$

Hence, the expression for the potential takes the form

$$\varphi = \frac{1}{4\pi\epsilon}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} \tag{14.3}$$

If instead of the two unlike charges of a dipole we have a more complicated neutral system of charges distributed in a volume V with density ρ, then the electrical properties of this system are described in the first approximation by the dipole moment $\mathbf{p}$ which is defined by

$$\mathbf{p} = \int_V \rho\mathbf{r}\,dV \tag{14.4}$$

where ρ is the volume charge density, and $\mathbf{r}$ is the radius vector of the point of integration. It is not difficult to see that the value of (14.4) is independent of the choice of origin. In fact, if the origin is moved to a point with radius vector $\mathbf{r}_0$ (Fig. 15), we then obtain

$$\mathbf{p}' = \int_V \rho\mathbf{r}'\,dV = \int_V \rho(\mathbf{r} - \mathbf{r}_0)\,dV$$

$$= \int_V \rho\mathbf{r}\,dV - \mathbf{r}_0\int_V \rho\,dV = \int_V \rho\mathbf{r}\,dV = \mathbf{p}$$

remembering that, since the system is neutral

$$\int \rho\,dV = 0$$

Thus, it has been shown that the definition (14.4) of the dipole moment is, in fact, independent of the choice of the origin of coordinates. If we use (14.4) to calculate the dipole moment of a dipole, then we obtain (14.1).

The electric field set up by a dipole equals

$$\mathbf{E} = -\operatorname{grad}\frac{1}{4\pi\epsilon}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} = \frac{1}{4\pi\epsilon}\left\{\frac{3(\mathbf{p}\cdot\mathbf{r})\,\mathbf{r}}{r^5} - \frac{\mathbf{p}}{r^3}\right\} \tag{14.5}$$

This equation shows that the intensity of the field of a dipole decreases in inverse proportion to the cube of the distance from the dipole, i.e., more rapidly than the field of a point charge.

Polarization of Dielectrics. A dielectric placed in an external electric field becomes *polarized*, i.e., it acquires a dipole moment distribution. The intensity of polarization is described by the polarization vector $\mathbf{P}$, which

is defined as the dipole moment per unit of volume. From this definition, it follows that the dipole moment $d\mathbf{p}$ of an element of volume dV of a dielectric is given by

$$d\mathbf{p} = \mathbf{P}\, dV \tag{14.6}$$

Phenomenological electrodynamics is not concerned with the molecular mechanism of polarization, which will be discussed in the electronic theory.

The magnitude of the polarization vector $\mathbf{P}$ at a given point is proportional to the electric field $\mathbf{E}$ at that point

$$\mathbf{P} = \kappa\epsilon_0\mathbf{E} \tag{14.7}$$

where the dimensionless parameter κ is called the electric susceptibility. It describes the ability of a dielectric to be polarized. In the absolute Gaussian system of units, the electric susceptibility κ' is 4π times smaller than κ in (14.7)

$$\kappa' = \frac{\kappa}{4\pi}$$

The susceptibility of the majority of good solid and liquid dielectrics lies between 1 and 10. The susceptibility of most gases is of the order of 0.0001 and almost always it may be ignored. However, there are dielectrics for which the susceptibility reaches considerably greater values. For water $\kappa = 80$, and for alcohol 25 to 30. There are some semiconductors which have susceptibilities of the order of 100,000.

Ferroelectrics are an important class of dielectrics (Rochelle salt, barium titanate, etc.). They are characterized by a nonlinear dependence of the polarization on the field intensity, and by the presence of residual polarization, i.e., the polarization does not disappear when the field which caused it disappears. In many ferroelectrics, the susceptibility is of the order of several thousands.

A substance can be polarized not only by an electric field; the polarization may also be produced by mechanical stresses. This is called the *piezoelectric effect.* It is observed in a number of substances, e.g., in quartz, and has a wide range of technical applications.

The polarization vector does not always have the same direction as the electric field vector. This anisotropy is most frequently observed in crystalline dielectrics. In such dielectrics, the polarization of the dielectric is different along different directions, and hence, it is impossible to describe the susceptibility by a single scalar κ. In this case, the susceptibility is described by an electric susceptibility tensor. Further discussion of these questions, however, lies outside the range of this book.

Scalar Potential in the Presence of a Dielectric. The effect of a dielectric

on an electrostatic field is to produce an additional field due to the polarization of the dielectric. An electric field in the presence of a dielectric is therefore the sum of two fields: (1) the field of free charges, i.e., charges not bound to the molecules and atoms of the dielectric, and (2) the field caused by the polarization of the dielectric.

Hence, the potential φ of the electric field may be written in the form

$$\varphi = \varphi_0 + \varphi_d \tag{14.8}$$

where φ_0 is the potential of the field of free charges, and φ_d is the potential of the field caused by the polarization of the dielectric.

Clearly

$$\varphi_0 = \frac{1}{4\pi\epsilon_0}\int_V \frac{\rho\, dV}{r} + \frac{1}{4\pi\epsilon_0}\int_S \frac{\sigma\, dS}{r} \tag{14.9}$$

where ρ and σ are the volume and surface densities of free charge.

From the equation for the potential of a dipole (14.3), it follows that the potential $d\varphi_d$ of a dipole of moment $d\mathbf{p}$ is given by

$$d\varphi_d = \frac{1}{4\pi\epsilon_0}\frac{d\mathbf{p}\cdot\mathbf{r}}{r^3} \tag{14.10}$$

Hence, taking (14.6) into account, it follows that

$$d\varphi_d = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{P}\cdot\mathbf{r}}{r^3}\, dV \tag{14.11}$$

and, therefore

$$\varphi_d = \frac{1}{4\pi\epsilon_0}\int_V \frac{\mathbf{P}\cdot\mathbf{r}}{r^3}\, dV \tag{14.12}$$

where the integration is carried out over the volume V of the dielectric.

This equation may be put in another form. In vector analysis, it is shown that

$$\operatorname{div}(\varphi\mathbf{A}) = \varphi \operatorname{div}\mathbf{A} + \mathbf{A}\cdot\operatorname{grad}\varphi \tag{14.13}$$

where φ is a scalar function, and $\mathbf{A}$ is a vector function of a point. We shall use this formula to transform the integrand in (14.12), remembering that here $\mathbf{r}$ is the radius vector from the element of volume dV to the point where the potential is measured. Hence, taking the operations div and grad as operations with respect to the coordinates of the element of volume dV, we have

$$\mathbf{P}\cdot\frac{\mathbf{r}}{r^3} = \mathbf{P}\cdot\operatorname{grad}\frac{1}{r} = \operatorname{div}\frac{\mathbf{P}}{r} - \frac{\operatorname{div}\mathbf{P}}{r} \tag{14.14}$$

Consequently

$$\varphi_d = \frac{1}{4\pi\epsilon_0}\int_V \frac{-\operatorname{div}\mathbf{P}}{r}\, dV + \frac{1}{4\pi\epsilon_0}\int_V \operatorname{div}\frac{\mathbf{P}}{r}\, dV \tag{14.15}$$

The second integral may be transformed in accordance with Gauss' theorem, which holds throughout the region where the integrand is continuous. But, by (14.7), the polarization vector **P** is discontinuous on the boundary separating different dielectrics. Therefore, Gauss' theorem may be applied to an arbitrary volume, provided that boundaries between different dielectrics are excluded from it. Let S be one such boundary (Fig. 16). We shall

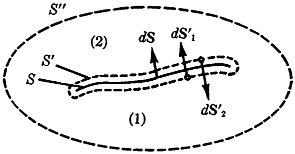


Fig. 16

exclude this surface from the region under consideration by an auxiliary surface S'. Then for the whole remaining volume, the integrand is continuous, and Gauss' theorem may be applied to it

$$\int_V \operatorname{div}\frac{\mathbf{P}}{r}\,dV = \int_{S''}\frac{\mathbf{P}}{r}\cdot d\mathbf{S} + \int_{S'}\frac{\mathbf{P}}{r}\cdot d\mathbf{S} \tag{14.16}$$

where S'' is the surface bounding the volume under consideration. We choose the positive normal to the surface so that it points in the direction of that dielectric which is denoted by the index 2 in Fig. 16. The other dielectric is denoted by the index 1. Then we have

$$\int_{S'}\frac{\mathbf{P}\cdot d\mathbf{S}}{r} = -\int_S\frac{\mathbf{P}_2\cdot d\mathbf{S}}{r} + \int_S\frac{\mathbf{P}_1\cdot d\mathbf{S}}{r} = \int_S\frac{P_{1n} - P_{2n}}{r}\,dS \tag{14.17}$$

The minus sign of the vector $\mathbf{P}_2$ arises as follows. When Gauss' theorem is applied to integration over the surface enclosing V, $d\mathbf{S}$ denotes an element of surface directed along the outward normal. Consequently, when we integrate over S' on the side of the medium 2, the vector $d\mathbf{S}'_2$ points in the direction opposite to the vector $d\mathbf{S}$ of the element of the boundary surface which points into the medium 2. Hence

$$\mathbf{P}_2\cdot d\mathbf{S}'_2 = -\mathbf{P}_2\cdot d\mathbf{S}$$

The quantity of $P_{1n} - P_{2n}$ describes the discontinuity of the normal components of the polarization vector at the boundary separating the dielectrics. Assuming that all dielectrics are contained in a finite region of space,

and taking S'' to be an infinitely distant surface, we see that on this surface $\mathbf{P} = 0$ and hence, the first integral in (14.16) vanishes.

Similar expressions are also obtained for the other boundaries separating dielectrics, and hence, if we take S to mean all the boundaries between dielectrics and V to mean all space, we may write

$$\varphi_d = \frac{1}{4\pi\epsilon_0}\int_V \frac{-\operatorname{div}\mathbf{P}}{r}\,dV + \frac{1}{4\pi\epsilon_0}\int_S \frac{P_{1n} - P_{2n}}{r}\,dS \tag{14.18}$$

If we compare equation (14.18) with equation (14.9) for the potential set up by free electric charges, it becomes clear that these equations are of analogous form, with ρ replaced by $-\operatorname{div}\mathbf{P}$ and σ by $P_{1n} - P_{2n}$. Hence, if we put

$$\rho_{\text{pol}} = -\operatorname{div}\mathbf{P} \qquad \sigma_{\text{pol}} = P_{1n} - P_{2n} \tag{14.19}$$

equation (14.18) may be written in the form

$$\varphi_d = \frac{1}{4\pi\epsilon_0}\int \frac{\rho_{\text{pol}}}{r}\,dV + \frac{1}{4\pi\epsilon_0}\int \frac{\sigma_{\text{pol}}}{r}\,dS \tag{14.20}$$

so that equation (14.8) for the total potential becomes, taking (14.9) and (14.20) into account

$$\varphi = \frac{1}{4\pi\epsilon_0}\int \frac{\rho + \rho_{\text{pol}}}{r}\,dV + \frac{1}{4\pi\epsilon_0}\int \frac{\sigma + \sigma_{\text{pol}}}{r}\,dS \tag{14.21}$$

The values ρ_{pol} and σ_{pol} are called, respectively, the *volume density* and the *surface density* of bound charges. This terminology is due to the fact that, in their role of setting up an additional electric field in the dielectric, these quantities are completely analogous to the volume and surface charge densities. But, as distinguished from the free charges, the bound charges cannot move freely in the dielectric.

From (14.20) it is evident that the bound surface charges appear on boundaries between different dielectrics and on the boundary between a dielectric and empty space, while the bound volume charges appear when the polarization is inhomogeneous, which may be due either to the inhomogeneity of the dielectric or to the inhomogeneity of the electric field.

Relationship Between the Susceptibility and the Permittivity. As has been shown, the presence of a dielectric may be completely described mathematically if, in addition to free charges, we also take into account bound charges. Thus, the electric field in a dielectric may be described by the same equations as the field in empty space, but with bound charges taken into account. Hence, the equations $\operatorname{div}\mathbf{E} = \rho/\epsilon_0$ must, if a dielectric is present, be written

$$\operatorname{div}\mathbf{E} = \frac{1}{\epsilon_0}(\rho + \rho_{\text{pol}}) \tag{14.22}$$

Using equation (14.19) for ρ_{pol} we obtain the expression

$$\operatorname{div} \mathbf{E} = \frac{1}{\epsilon_0} (\rho - \operatorname{div} \mathbf{P})$$

which may be rewritten, for convenience

$$\operatorname{div} (\epsilon_0 \mathbf{E} + \mathbf{P}) = \rho \tag{14.23}$$

On the other hand, we know Maxwell's equation for the electric induction vector **D** in a dielectric (2.17). But equations (14.23) and (2.17) describe the same field in the dielectric, and hence, it follows from these equations that

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{14.24}$$

Taking into account the relationships $\mathbf{D} = \epsilon \mathbf{E}$, $\mathbf{P} = \kappa \epsilon_0 \mathbf{E}$, equation (14.24) may be rewritten in the form

$$\epsilon \mathbf{E} = \epsilon_0 \mathbf{E} + \kappa \epsilon_0 \mathbf{E}$$

whence it follows that

$$\epsilon = \epsilon_0 (1 + \kappa) \qquad \kappa = \frac{\epsilon}{\epsilon_0} - 1 \tag{14.25}$$

Since $\epsilon > \epsilon_0$ always, κ is always positive.

In the absolute Gaussian system of units, the permittivity ϵ' and the susceptibility κ' are related to ϵ and κ by the relationships

$$\epsilon = \epsilon' \epsilon_0 \qquad \kappa = 4\pi\kappa'$$

Hence, in the Gaussian system of units, equation (14.25) takes the form

$$\epsilon' = 1 + 4\pi\kappa' \qquad \kappa' = \frac{\epsilon' - 1}{4\pi}$$

Method of Images. We apply the method of images to dielectrics in a manner similar to the case of conductors described in the preceding section.

Let us consider two semi-infinite media with a plane surface of separation between them (Fig. 17). The permittivities of the first and second media are, respectively, ϵ_1 and ϵ_2. Let there be a point charge q in the first medium

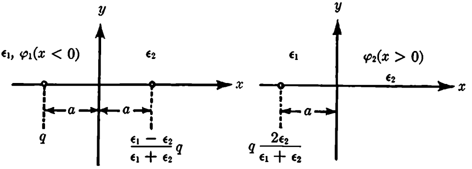


Fig. 17

at a distance a from the boundary. It is asserted that the potential in the first medium will be the same as that due to q together with an image charge $[(\epsilon_1 - \epsilon_2)q]/(\epsilon_1 + \epsilon_2)$ in the second medium at a distance a from the boundary. At this stage, the permittivity of both media is taken to be equal to ϵ_1. The potential in the second medium will be equal to the potential due to a charge $\dfrac{2\epsilon_2 q}{\epsilon_1 + \epsilon_2}$ at the site of the charge q in the first medium. At this stage, the permittivity of the whole medium is taken to be equal to ϵ_2. Thus, the potentials φ_1 and φ_2 in the first and second media, respectively, are given by the expressions

$$\varphi_1 = \frac{q}{4\pi\epsilon_1}\left[\frac{1}{\sqrt{(x+a)^2 + y^2}} + \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}\frac{1}{\sqrt{(x-a)^2 + y^2}}\right]$$

$$\varphi_2 = \frac{q}{4\pi\epsilon_2}\frac{2\epsilon_2}{\epsilon_1 + \epsilon_2}\frac{1}{\sqrt{(x+a)^2 + y^2}}$$

It is not difficult to show, by direct differentiation, that φ_1 and φ_2 satisfy Laplace's equation and the necessary boundary conditions

$$\epsilon_1 \left.\frac{\partial\varphi_1}{\partial x}\right|_{x=0} = \epsilon_2 \left.\frac{\partial\varphi_2}{\partial x}\right|_{x=0} \qquad \left.\frac{\partial\varphi_1}{\partial y}\right|_{x=0} = \left.\frac{\partial\varphi_2}{\partial y}\right|_{x=0}$$

$$\varphi_1\Big|_{x\to-\infty} \longrightarrow 0 \qquad \varphi_2\Big|_{x\to\infty} \longrightarrow 0$$

The force acting on the charge q is equal to the force between the charge and its image $[(\epsilon_1 - \epsilon_2)/(\epsilon_1 + \epsilon_2)]q$ at a distance $2a$ from q, i.e., it is equal to

$$F = \frac{1}{4\pi\epsilon_1}\left(\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}\right)\frac{q^2}{4a^2}$$

If $\epsilon_1 < \epsilon_2$, F is negative. This means that in this case q is attracted toward the boundary between the dielectrics. If $\epsilon_1 > \epsilon_2$, then F is positive and q is therefore repelled from the boundary.

In other cases the method is applied in a similar manner: it is necessary to find a system of image charges such that the resultant potential will satisfy Laplace's equation and the boundary conditions on the surface of separation between the dielectrics.

§15. Energy of the Electrostatic Field and the Energy of the Interaction Between Charges

Expression for the Energy in Terms of the Field Vector. From the general formula (8.12) for the energy of the electromagnetic field we obtain the expression for the electrostatic field

$$W = \frac{1}{2}\int_V \mathbf{D}\cdot\mathbf{E}\,dV \tag{15.1}$$

Expression for the Energy in Terms of the Potential and Charge Density. We replace $\mathbf{E}$ in (15.1) by the scalar potential $\mathbf{E} = -\text{grad}\,\varphi$, and use the formula from vector analysis

$$-\mathbf{D}\cdot\text{grad}\,\varphi = \varphi\,\text{div}\,\mathbf{D} - \text{div}\,(\varphi\mathbf{D})$$

Then (15.1) becomes

$$W = \frac{1}{2}\int \varphi\rho\,dV - \frac{1}{2}\int \text{div}\,(\varphi\mathbf{D})\,dV \tag{15.2}$$

because $\text{div}\,\mathbf{D} = \rho$. The second integral in (15.2) may be transformed by Gauss' formula into a surface integral. The potential φ is continuous at all points of space, but $\mathbf{D}$ is discontinuous on charged surfaces. Hence, these charged surfaces must be taken out of the region of integration in exactly the same way as in § 14 (Fig. 16). Repeating the arguments of §14, we obtain the expressions

$$\int_V \text{div}\,(\varphi\mathbf{D})\,dV = \int_{S''} \varphi\mathbf{D}\cdot d\mathbf{S} + \int_{S'} \varphi\mathbf{D}\cdot d\mathbf{S} \tag{15.3}$$

$$\int_{S'} \varphi\mathbf{D}\cdot d\mathbf{S} = -\int_S \varphi\mathbf{D}_2\cdot d\mathbf{S} + \int_S \varphi\mathbf{D}_1\cdot d\mathbf{S} = \int_S \varphi(D_{1n} - D_{2n})\,dS \tag{15.4}$$

Taking the boundary condition (9.15) into account, $D_{2n} - D_{1n} = \sigma$, we shall write (15.4) in the form

$$\int_{S'} \varphi\mathbf{D}\cdot d\mathbf{S} = -\int_S \varphi\sigma\,dS$$

If the surface S'' in (15.3) is removed to infinity, and it is assumed that all the charges are in a finite region of space, then the integral over S'' tends to zero. This is clear from the following estimates. The potential of a point charge decreases with distance as $1/r$, where r is the distance from some point of the system. The field $\mathbf{E}$, being the gradient of the potential, decreases at great distances from the system of charges as $1/r^2$. Consequently, the integrand $\varphi\mathbf{D}$ in this integral decreases as $1/r^3$. The area of integration increases as r^2. Hence, it follows that when the surface of integration S'' moves away, the integral decreases as $1/r$, i.e., in the limit $r \to \infty$ the integral tends to zero. Hence, (15.2) finally takes the form

$$W = \frac{1}{2}\int_V \rho\varphi\,dV + \frac{1}{2}\int_S \sigma\varphi\,dS \tag{15.5}$$

In this integral, V means all space and S means all charged surfaces in space.

Numerically, this expression gives the same result as (15.1), but its physical content is somewhat different. Equation (15.1) expresses the fact that the energy of the electrostatic field is distributed throughout all space with the density

$$U_e = \frac{1}{2}\mathbf{D}\cdot\mathbf{E} \tag{15.6}$$

Equation (15.5) expresses the fact that the energy of the electromagnetic field is the energy of the interaction between charges: an element of charge $\rho\, dV$ in a potential field φ set up by other charges has the potential energy $\varphi\rho\, dV$. The factor $\frac{1}{2}$ in the integral appears because, in the course of integration, every element of charge is counted twice: once to calculate the potential energy of a given charge in the field of the other charges, and once to calculate the potential energy of the other charges in the field of the given charge.

Energy of the Interaction Between Point Charges. A point charge may be considered as the limiting case of a charge distributed over a small region of space, when the dimensions of the region tend to zero, but the amount of charge contained in the region remains constant. In this case, the charge density at the point to which the region shrinks tends to infinity, while at other points of space the charge density becomes zero. Hence, the volume charge density when there is a point charge e at a point $\mathbf{r}_0$ may be written as a δ-function

$$\rho(\mathbf{r}) = e\,\delta(\mathbf{r} - \mathbf{r}_0) \tag{15.7}$$

when the δ-function (delta-function) is defined as follows: $\delta(\mathbf{r})$ is equal to zero at all points excluding $\mathbf{r} = 0$. At $\mathbf{r} = 0$ this function tends to infinity in such a way that the integral of this function over a region containing the point $\mathbf{r} = 0$ is equal to 1. Thus, we may write

$$\delta(\mathbf{r}) = \begin{cases} 0 & \text{for } \mathbf{r} \neq 0 \\ \infty & \text{for } \mathbf{r} = 0 \end{cases} \tag{15.7a}$$

$$\int_V \delta(\mathbf{r})\, dV = \begin{cases} 0 & \text{if the point } \mathbf{r} = 0 \text{ lies outside the region of integration } V. \\ 1 & \text{if the point } \mathbf{r} = 0 \text{ lies inside the region of integration } V. \end{cases} \tag{15.7b}$$

From this property of the δ-function, we obtain a rule for integration when such a function is included in the integrand

$$\int_V f(\mathbf{r})\ \delta(\mathbf{r} - \mathbf{r}_0)\, dV = f(\mathbf{r}_0) \tag{15.8}$$

The point $\mathbf{r}_0$ lies within the region of integration V.

When several point charges are present, we may consider them as the

limiting case of charges of finite size occupying infinitely small volumes ΔV_i, where $\Delta V_i \to 0$. Remembering that the surface integrals become zero in this case, we can rewrite (15.5) in the form

$$W = \frac{1}{2} \sum_i \int_{\Delta V_i} \rho_i \varphi_i \, dV_i \tag{15.9}$$

where ΔV_i is the volume occupied by the i^{th} charge, φ_i is the potential set up both by other charges and by the i^{th} charge itself. Thus, we may write

$$\varphi_i = \varphi'_i + \varphi_i^{\text{self}}$$

where φ'_i is the potential created by all charges except the i^{th}, φ_i^{self} is the potential created by the elements of the i^{th} charge itself. Hence, (15.9) may be put in the form

$$W = \frac{1}{2} \sum_i \int_{\Delta V_i} \rho_i \varphi'_i \, dV_i + \frac{1}{2} \sum_i \int_{\Delta V_i} \rho_i \varphi_i^{\text{self}} \, dV_i \tag{15.10}$$

In the limit $\Delta V_i \to 0$, we obtain, for the integrals in the first sum

$$\lim_{\Delta V_i \to 0} \int_{\Delta V_i} \rho_i \varphi'_i \, dV_i = e_i \varphi_i$$

since in this case $\rho_i = e_i \delta(\mathbf{r} - \mathbf{r}_i)$, and we may use (15.8) for the integration. The integrals in the second sum in (15.10) give infinity, as $\Delta V_i \to 0$, since not only does $\rho_i \to \infty$, but also $\varphi_i^{\text{self}} \to \infty$. Hence, after proceeding to the limit $\Delta V_i \to 0$, equation (15.10) takes the form

$$W = \frac{1}{2} \sum_i e_i \varphi'_i + \infty$$

The second infinite term is equal to the infinite energy between the elements of the same point charge. This term is independent of the mutual positions of the point charges and may be dropped, so that the energy of the interaction between charges may be put in the form

$$W = \frac{1}{2} \sum e_i \varphi_i \tag{15.11}$$

where φ_i is the potential at the site of the charge e_i caused by all charges except the i^{th}. For simplicity, we shall write φ_i without the prime.

The energy of a point charge in a potential field φ is equal to

$$W = e\varphi \tag{15.12}$$

The coefficient $\frac{1}{2}$ in the expression (15.11) arises due to the fact that in the summation each point charge is considered twice: once when its energy in the field of all other charges is considered, and a second time when the energy of the other charges in its field is considered.

It must be stressed once again, that in passing from (15.5) to (15.11) we drop the term for the infinite energy of interaction between different elements of the same point charge, which is usually called the *self-energy*. The value of (15.5) is numerically equal to (15.1), which is always positive, since

$$\mathbf{D}\cdot\mathbf{E} = \epsilon \mathbf{E}^2 > 0 \tag{15.13}$$

The value of (15.11), however, as is not difficult to see, may be negative. To verify this, it is sufficient to consider the energy of two unlike charges.

Hence, we can say that the total energy of the electromagnetic field due to charges distributed in space with a finite density is given by (15.5) and (15.1) and is positive everywhere. In the case of point charges, the energy remains positive, but may become infinitely great. If we leave out the infinite self-energy of point charges, then the energy of interaction between the charges, given by (15.11) remains finite, but may be either positive or negative.

Equation (15.11) may be transformed, with the aid of (11.3), for a potential due to a system of point charges, which may be written

$$\varphi_i = \frac{1}{4\pi\epsilon}\sum_{j\neq i}\frac{e_j}{r_{ij}} \tag{15.14}$$

where φ_i is the potential at the site of the i^{th} charge due to the other charges, $j \neq i$; r_{ij} is the distance between the i^{th} and the j^{th} point charges. Substituting (15.14) in (15.11) we find

$$W = \frac{1}{8\pi\epsilon}\sum_{i\neq j}\frac{e_i e_j}{r_{ij}} \tag{15.15}$$

We write this equation for the special case of two charges

$$W = \frac{1}{8\pi\epsilon}\left(\frac{e_1e_2}{r_{12}} + \frac{e_2e_1}{r_{21}}\right) = \frac{1}{4\pi\epsilon}\frac{e_1e_2}{r_{12}} = e_1\varphi_1 = e_2\varphi_2 \tag{15.16}$$

where φ_1 is the potential due to the charge e_2 at the site of e_1. This is the energy of interaction between point charges e_1 and e_2 at a distance r_{12} apart. For unlike charges, this energy is negative.

Energy of Charged Conductors. In conductors there are no volume charges, hence, $\rho = 0$ in equation (15.5). The potential φ is constant on each conductor. Hence, in the case of conductors, (15.5) takes the form

$$W = \frac{1}{2}\sum_i\int_{S_i}\varphi_i\sigma_i\,dS = \frac{1}{2}\sum_i\varphi_i\int_{S_i}\sigma_i\,dS = \frac{1}{2}\sum_i\varphi_i q_i \tag{15.17}$$

where φ_i and σ_i are the potential and surface density of the i^{th} conductor, and q_i is its charge.

Using this expression to evaluate the energy of a charged capacitor, and taking (13.16) into account, we obtain

$$W = \frac{1}{2}q(\varphi_1 - \varphi_2) = \frac{1}{2}qV = \frac{1}{2}CV^2 = \frac{q^2}{2C} \tag{15.18}$$

where q, V, and C denote, respectively, the charge on the capacitor plate, the potential difference between the plates, and the capacitance.

Energy of a Dipole in an External Field. The energy of a dipole (14.1) in an external electric field is the sum of the energies of its charges. Hence

$$W = q[\varphi(\mathbf{r} + \mathbf{l}) - \varphi(\mathbf{r})] \tag{15.19}$$

Expanding the function $\varphi(\mathbf{r} + \mathbf{l})$ as a Taylor series

$$\varphi(\mathbf{r} + \mathbf{l}) = \varphi(\mathbf{r}) + l_x\frac{\partial \varphi}{\partial x} + l_y\frac{\partial \varphi}{\partial y} + l_z\frac{\partial \varphi}{\partial z} + \cdots$$

and taking only first-order terms in $\mathbf{l}$, we find

$$\varphi(\mathbf{r} + \mathbf{l}) - \varphi(\mathbf{r}) = -l_xE_x - l_yE_y - l_zE_z$$

where $\mathbf{E} = -\text{grad}\,\varphi$. Thus, finally, we obtain the following expression for the energy of a dipole in an external electric field $\mathbf{E}$

$$W = -\mathbf{p}\cdot\mathbf{E} \tag{15.20}$$

where $\mathbf{p}$ is the moment of the dipole.

§16. Mechanical Energy in an Electrostatic Field

Force on a Point Charge. The force acting on a point charge placed in an external electric field $\mathbf{E}$ is

$$\mathbf{F} = e\mathbf{E} \tag{16.1}$$

Force on a Continuous Charge Distribution. If the charge is distributed in a volume V with a volume density ρ, then the force acting upon it in an electric field $\mathbf{E}$ is

$$\mathbf{F} = \int \rho\mathbf{E}\,dV \tag{16.2}$$

Force on a Dipole. The force on a dipole is the sum of the forces acting on the charges of the dipole (Fig. 18)

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 = e(\mathbf{E}_2 - \mathbf{E}_1) \tag{16.3}$$

where $\mathbf{E}_2$ is the electric field intensity at the site of the positive charge of the dipole, and $\mathbf{E}_1$ is the field intensity at the site of the negative charge. Assuming that the distance l between the charges is small compared with

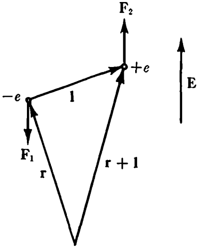


Fig. 18

the distances at which the external field changes significantly, we may expand the field $\mathbf{E}_2$ in a Taylor series and take only the first term

$$\mathbf{E}_2 \equiv \mathbf{E}\ (\mathbf{r} + \mathbf{l}) = \mathbf{E}\ (\mathbf{r}) + l_x \frac{\partial \mathbf{E}}{\partial x} + l_y \frac{\partial \mathbf{E}}{\partial y} + l_z \frac{\partial \mathbf{E}}{\partial z} + \cdots = \mathbf{E}_1 + (\mathbf{l} \cdot \nabla)\ \mathbf{E} + \cdots \quad (16.4)$$

where we introduce the notation

$$(\mathbf{l} \cdot \nabla)\ \mathbf{E} = \left(l_x \frac{\partial}{\partial x} + l_y \frac{\partial}{\partial y} + l_z \frac{\partial}{\partial z} \right) \mathbf{E} = l_x \frac{\partial \mathbf{E}}{\partial x} + l_y \frac{\partial \mathbf{E}}{\partial y} + l_z \frac{\partial \mathbf{E}}{\partial z}$$

and use the fact that $\mathbf{E}(\mathbf{r}) = \mathbf{E}_1$ is the field intensity at the site of the negative charge.

Taking (16.3) and (16.4) into account, and that $\mathbf{p} = e\mathbf{l}$, we find the following expression for the force on the dipole

$$\mathbf{F} = (\mathbf{p} \cdot \nabla)\ \mathbf{E} \quad (16.5)$$

In a homogeneous electric field, this force is equal to zero, since the forces on the unlike charges of the dipole are equal and opposite, and thus, cancel each other out. The greater the inhomogeneity of the electric field, the stronger the force on the dipole.

Moment of Forces Acting on a Dipole. From Fig. 18 it is clear that there is a couple acting on the dipole, and that its moment about the center of the dipole is equal to

$$\mathbf{N} = \mathbf{l} \times e\mathbf{E} = \mathbf{p} \times \mathbf{E} \quad (16.6)$$

This moment tends to turn the dipole moment to coincide with the direction of the field $\mathbf{E}$.

Forces on a Conductor. On the outer surface of a conductor, the field is equal to

$$\mathbf{E} = \frac{\sigma}{\epsilon}\mathbf{n} \tag{16.7}$$

where $\mathbf{n}$ is the unit vector along the outward normal and σ is the surface charge density. We consider an infinitely small element of surface dS, and the field close to this element (Fig. 19). This field is due to both the charges

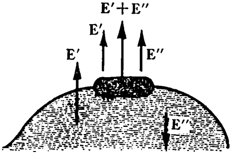


Fig. 19

on the surface element dS, and all the other charges. It is clear that the only field acting on the charges of dS is the field set up by all the other charges outside dS. We denote this field by $\mathbf{E}'$. The field due to the charges on dS we denote by $\mathbf{E}''$. The field $\mathbf{E}'$ has the same magnitude and direction on both sides of dS, i.e., it is a continuous field. The field $\mathbf{E}''$, however, has the same magnitude, but opposite directions on opposite sides of the surface. Outside the conductor, $\mathbf{E}'$ and $\mathbf{E}''$ combine to give the resultant field $\mathbf{E}$ defined by (16.7)

$$\mathbf{E}' + \mathbf{E}'' = \mathbf{E} = \frac{\sigma}{\epsilon}\mathbf{n} \tag{16.8}$$

Inside the conductor, the total field is zero. Hence, the two fields acting in opposite directions must be equal in magnitude

$$E' = E'' \tag{16.9}$$

Hence, it follows from (16.8) that

$$E' = E'' = \frac{1}{2}\frac{\sigma}{\epsilon} \tag{16.10}$$

and, in particular

$$\mathbf{E}' = \frac{1}{2}\frac{\sigma}{\epsilon}\mathbf{n} \tag{16.11}$$

The charge on dS is $dq = \sigma\, dS$, and the field $\mathbf{E}'$ acts on this charge. Consequently, the force on this charge is

$$d\mathbf{F} = dq\mathbf{E}' = \frac{1}{2}\frac{\sigma^2}{\epsilon}\mathbf{n}\, dS \tag{16.12}$$

Hence, it follows that the surface density of the force $\mathbf{f}_n$, i.e., the force acting on unit surface of the conductor, is equal to

$$\mathbf{f}_n = \frac{d\mathbf{F}}{dS} = \frac{1}{2}\frac{\sigma^2}{\epsilon}\mathbf{n} \tag{16.13}$$

Eliminating the surface charge density σ by means of (16.7), we obtain the alternative expression for the surface density of the force

$$\mathbf{f}_n = \frac{1}{2}\epsilon E^2\mathbf{n} \tag{16.14}$$

where $\mathbf{E}$ is the total field on the surface of the conductor.

Equations (16.13) and (16.14) show that there is always a force acting on the surface of a conductor along the outward normal to the surface. The density of this force is numerically equal to the density of the electrostatic field energy on the surface of the conductor.

To evaluate the total force on a conductor in an electrostatic field, we must integrate the density of the force over the whole surface S of the conductor

$$\mathbf{F} = \int_S \mathbf{f}_n \, dS = \frac{\epsilon}{2}\int_S E^2 \, d\mathbf{S} = \frac{1}{2\epsilon}\int_S \sigma^2 \, d\mathbf{S} \tag{16.15}$$

In (16.15), $d\mathbf{S}$ denotes the vector of the surface element of the conductor directed along the outward normal, $d\mathbf{S} = \mathbf{n}\, dS$.

Forces on a Dielectric. Since the dipole moment of an element of volume dV of a polarized dielectric in an electrostatic field $\mathbf{E}$ is equal to $d\mathbf{p} = \mathbf{P}\, dV$, the force acting on this element will be, according to (16.15)

$$d\mathbf{F} = (d\mathbf{p}\cdot\nabla)\,\mathbf{E} = (\mathbf{P}\cdot\nabla)\,\mathbf{E}\, dV \tag{16.16}$$

Hence, for a volume density of the force $\mathbf{f}$, i.e., the force acting on unit volume of the dielectric, we obtain the expression

$$\mathbf{f} = \frac{d\mathbf{F}}{dV} = (\mathbf{P}\cdot\nabla)\,\mathbf{E} \tag{16.17}$$

Taking (14.25) and (14.7) into account, this expression may be rewritten

$$\mathbf{f} = (\epsilon - \epsilon_0)\,(\mathbf{E}\cdot\nabla)\,\mathbf{E} \tag{16.18}$$

Using formula (A.10) in Appendix 1 with $\mathbf{A} = \mathbf{B} = \mathbf{E}$

$$(\mathbf{E}\cdot\nabla)\,\mathbf{E} = \frac{1}{2}\operatorname{grad} E^2 - \mathbf{E}\times\operatorname{curl}\mathbf{E} \tag{16.19}$$

and $\operatorname{curl}\mathbf{E} = 0$ for the electrostatic field, we may rewrite the expression for the volume density of the force as

$$\mathbf{f} = \frac{\epsilon - \epsilon_0}{2}\operatorname{grad} E^2 = \frac{\kappa\epsilon_0}{2}\operatorname{grad} E^2 \tag{16.20}$$

It is clear from this expression that the direction of the force on the dielectric is independent of the direction of the field, but always lies in the direction of increase of the absolute magnitude of the field. This means that the dielectric is always attracted toward the region of greatest absolute magnitude of the electrostatic field.

Evaluation of Forces from the Expression for the Energy. Let us suppose that the energy W of a system depends on several parameters $q_1, q_2, \ldots, q_n$

$$W = W(q_1, q_2, \ldots, q_n) \tag{16.21}$$

When these parameters change by an amount δq_i, the energy of the system receives an increment

$$\delta W = \sum_{i=1}^{n} \frac{\partial W}{\partial q_i} \delta q_i \tag{16.22}$$

Let the work done be δA. Then, by the law of conservation of energy, this is equal to the change in the energy of the system with the sign reversed

$$\delta A = -\delta W \tag{16.23}$$

But this work is equal to the sum of the work done by the generalized forces F_i, related to the change in the parameters q_i, i.e.

$$\delta A = \sum_{i=1}^{n} F_i \, \delta q_i \tag{16.24}$$

Taking (16.24) and (16.22) into account, (16.23) becomes

$$\sum_{i=1}^{n} F_i \, \delta q_i = -\sum_{i=1}^{n} \frac{\partial W}{\partial q_i} \delta q_i \tag{16.25}$$

Remembering that the parameters q_i are independent, and using (16.25), we can write down the following expression for the generalized forces acting on the system under consideration

$$F_i = -\frac{\partial W}{\partial q_i} \tag{16.26}$$

If the parameter q_i is an ordinary space coordinate, then F_i is the corresponding component of the ordinary force.

Thus, the problem of finding the forces acting on a system reduces to the problem of finding the energy of the system as a function of the parameters which describe its state. The derivatives of the energy with respect to the parameters give, with signs reversed, the generalized forces acting on the system. To evaluate the derivative (16.26), it is, of course, necessary to observe which values in it remain constant.

As an example, let us consider the forces acting on a dipole. The energy of the dipole is given by (15.20). Hence, the force on the dipole equals

$$\mathbf{F} = \operatorname{grad} \mathbf{p} \cdot \mathbf{E} \tag{16.27}$$

We use the vector analysis formula (A.10) of Appendix 1 for the gradient of a scalar product, substituting **p** for **A**, and **E** for **B**. Bearing in mind that the dipole moment **p** does not depend explicitly on the coordinates, and taking into account the potential nature of the electrostatic field (curl $\mathbf{E} = 0$), we transform (16.27) into the form

$$\mathbf{F} = (\mathbf{p}\cdot\nabla)\,\mathbf{E} \tag{16.28}$$

This expression is the same as (16.5), which has been deduced by a different method.

Sometimes it is convenient to put (16.28) in the form

$$F_x = \mathbf{p}\cdot\frac{\partial \mathbf{E}}{\partial x} \qquad F_y = \mathbf{p}\cdot\frac{\partial \mathbf{E}}{\partial y} \qquad F_z = \mathbf{p}\cdot\frac{\partial \mathbf{E}}{\partial z}$$

which, for curl $\mathbf{E} = 0$, is equivalent to (16.28).

PROBLEMS

1 Using Gauss' theorem, find the field set up by an infinite straight filament, uniformly charged with a line density τ (Fig. 20). Evaluate the field, if $\tau = (\frac{1}{9})10^{-9}$ coul/m, $d = 20$ cm.

Solution: From considerations of symmetry, it is clear that the electric field vector lies in planes perpendicular to the filament. We construct a circular cylinder of radius d, the axis of which coincides with the filament. Let the height of the cylinder be l. Applying Gauss' theorem to the cylinder, we have

$$\int_S \mathbf{D}\cdot d\mathbf{S} = q$$

where q is the total charge in the cylinder, S is the surface of the cylinder. It is evident that $q = \tau l$. The flux of **D** through the end surfaces of the cylinder is equal to zero, since **D** is parallel to the end surfaces. The flux of **D** through the lateral surface is easy to calculate, since at every point on this surface the direction of **D** coincides with the direction of $d\mathbf{S}$, while the magnitude of **D** is constant. Hence, we have

$$\int_S \mathbf{D}\cdot d\mathbf{S} = D\,2\pi\,dl$$

Thus, Gauss' theorem reduces to the equation

$$\epsilon_0 E\,2\pi\,dl = \tau l$$

Whence it follows that

$$E = \frac{1}{2\epsilon_0}\frac{\tau}{\pi d} = 10\ \mathrm{V/m}$$

Comparing this expression with equation (12.21) for the field outside an infinite charged cylinder, we see that the fields of an infinite charged cylinder

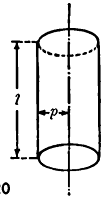


Fig. 20

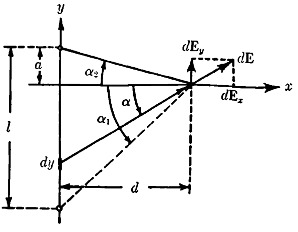


Fig. 21

and an infinite charged filament are the same if they carry the same amount of electricity per unit length (for the cylinder, $\tau = \rho\pi a^2$).

2 Find the field of a straight filament of finite length l, uniformly charged with a line density τ (Fig. 21). Consider the numerical example: $\tau = 10^{-10}$ coul/m, $l = 1$ m, $d = 50$ cm, $a = 50$ cm.

Solution: We use Coulomb's law

$$dE_x = \frac{\tau\, dy}{4\pi\epsilon_0(y^2 + d^2)} \cos\alpha = \frac{\tau\, dy}{4\pi\epsilon_0(y^2 + d^2)} \frac{d}{\sqrt{y^2 + d^2}}$$

$$dE_y = \frac{\tau\, dy}{4\pi\epsilon_0(y^2 + d^2)} \sin\alpha = \frac{\tau\, dy}{4\pi\epsilon_0(y^2 + d^2)} \left(\frac{-y}{\sqrt{y^2 + d^2}}\right)$$

Hence, it follows that

$$E_x = \frac{\tau d}{4\pi\epsilon_0} \int_{-(l-a)}^{a} \frac{dy}{(y^2 + d^2)^{3/2}} \qquad E_y = -\frac{\tau}{4\pi\epsilon_0} \int_{-(l-a)}^{a} \frac{y\, dy}{(y^2 + d^2)^{3/2}}$$

Altering the variables

$$y = d\tan\alpha \qquad dy = \frac{d}{\cos^2\alpha}\, d\alpha \qquad 1 + \tan^2\alpha = \frac{1}{\cos^2\alpha}$$

we obtain

$$E_x = \frac{\tau}{4\pi\epsilon_0 d}(\sin\alpha_2 + \sin\alpha_1) \approx 1.27 \text{ V/m}$$

$$E_y = \frac{\tau}{4\pi\epsilon_0 d}(\cos\alpha_2 - \cos\alpha_1) = 0$$

If the length of the filament is infinite ($l \to \infty$) then $\alpha_1 = \alpha_2 = \pi/2$, $E_y = 0$ and E_x is identical with the expression obtained in the preceding example.

3 A ring of radius a is uniformly charged. The total charge of the ring is Q. Determine the field intensity and the potential at points on the axis of the ring (Fig. 22).

Consider the numerical example: $a = 5$ cm, $h = 3$ cm, $Q = 2 \times 10^{-13}$ coul.

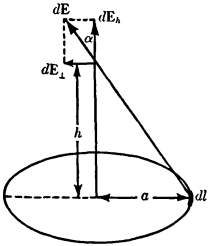


Fig. 22

Solution: We determine the field intensity directly from Coulomb's law. The amount of electricity on a line element of length dl is $dq = (Q/2\pi a)\,dl$. Hence, we have

$$dE_h = \frac{Q\,dl\cos\alpha}{4\pi\epsilon_0\,2\pi a(a^2+h^2)} = \frac{Q\,dlh}{4\pi\epsilon_0\,2\pi a(a^2+h^2)^{3/2}}$$

whence, integrating along the ring, we find

$$E_h = \frac{Qh}{4\pi\epsilon_0(a^2+h^2)^{3/2}} \approx 0.25\text{ V/m} \tag{1}$$

The components of the field intensity of the various elements of the ring perpendicular to the axis of the ring add up to zero, as may be seen directly from Fig. 22.

The potential may be calculated from equation (11.10), taking as the path of integration the axial direction, the field along which is known and is given by equation (1) above

$$\varphi(h) = \int_h^\infty \frac{Qh\,dh}{4\pi\epsilon_0(a^2+h^2)^{3/2}} = \frac{1}{4\pi\epsilon_0}\frac{Q}{\sqrt{a^2+h^2}} = 1.5\times 10^{-2}\text{ V}$$

4 Determine the intensity of the field on the axis of a disc if a charge Q is distributed uniformly on the disc. The radius of the disc is a (Fig. 23).

Consider the numerical example: $Q = 10^{-10}$ coul, $a = 10$ cm, $h = 20$ cm.

Solution: Using equation (11.15), we have

$$\varphi(h) = \frac{1}{4\pi\epsilon_0}\int\frac{\sigma\,dx\,dy}{\sqrt{x^2+y^2+h^2}} \qquad \sigma = \frac{Q}{\pi a^2}$$

The integral may be conveniently evaluated in polar coordinates

$$x^2+y^2 = r^2 \qquad dx\,dy = dS = r\,dr\,d\alpha$$

$$\varphi(h) = \frac{\sigma}{4\pi\epsilon_0}\int_0^{2\pi} d\alpha\int_0^a \frac{r\,dr}{\sqrt{r^2+h^2}} = \frac{1}{2\pi\epsilon_0}\frac{Q}{a^2}\left(\sqrt{a^2+h^2}-h\right) \approx 4.3\text{ V}$$

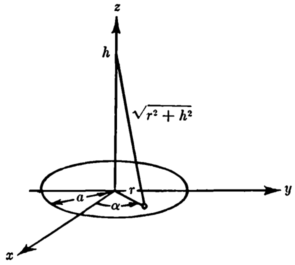


Fig. 23

whence

$$E_h = -\frac{\partial \varphi}{\partial h} = \frac{1}{2\pi\epsilon_0}\frac{Q}{a^2}\left(1 - \frac{h}{\sqrt{a^2 + h^2}}\right) \approx 18 \text{ V/m}$$

5 Find the field inside and outside a uniformly charged sphere of radius a by solving Poisson's equation and using Gauss' theorem. The volume charge density is ρ. The potential is normalized by equating to zero at infinity. *Hint:* Use the expression for Laplace's operator in spherical coordinates, and carry out the calculation in a similar manner to that used to deduce equation (12.21). Use Gauss' theorem in a manner similar to that of example 1 of this chapter, calculating the flux of **D** over a sphere concentric with the given sphere.

Answer:
$$\begin{cases} \varphi(r) = \dfrac{Q}{4\pi\epsilon_0 a} + \dfrac{\rho}{6\epsilon_0}(a^2 - r^2) \\ \mathbf{E}(r) = \dfrac{\rho}{3\epsilon_0}\mathbf{r} \quad \text{for} \quad 0 < r < a \end{cases}$$

$$\begin{cases} \varphi(r) = \dfrac{1}{4\pi\epsilon_0}\dfrac{Q}{r} \\ \mathbf{E}(r) = \dfrac{Q}{4\pi\epsilon_0 r^2}\dfrac{\mathbf{r}}{r} \quad \text{for} \quad a < r \end{cases}$$

Q is the total charge on the sphere: $Q = \dfrac{4\pi}{3}a^3\rho$.

6 Find the field in an infinite cylindrical cavity in an infinite uniformly charged circular cylinder. The volume density is ρ, and the other values are given in Fig. 24.
Solution: From the point of view of the effect on the electric field, the presence of a cavity is equivalent to the presence, in a solid cylinder, of charges of

opposite sign, uniformly filling the cavity with a density ρ. Hence, the field due to a cylinder with a cavity may be considered as the superposition of two fields: the field of a solid cylinder of the larger radius, with a charge density ρ, and the field of a solid cylinder filling the cavity, with a charge density $-\rho$. Using equation (12.21), we obtain

$$\mathbf{E} = \frac{1}{2}\frac{\rho}{\epsilon_0}(\mathbf{R} - \mathbf{r}) = \frac{1}{2}\frac{\rho}{\epsilon_0}\mathbf{r}_0$$

i.e., the field in the cavity is homogeneous. The field intensity vector is directed along the line joining the axis of the cylinder and the axis of the cavity.

Using a completely analogous method, we may find the field at all points of space outside the cavity.

From the above formula it follows that when the axis of the cavity and the axis of the cylinder coincide, the field in the cavity is equal to zero.

7 Find the electric field in a spherical cavity inside a uniformly charged sphere (Fig. 24).

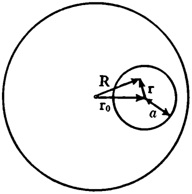


Fig. 24

Hint: Use calculations similar to those of the solution of Problem 6.

Answer: $\mathbf{E} = \frac{\rho}{3\epsilon_0}\mathbf{r}_0$

8 Find the capacitance of a conducting sphere of radius r.
Solution: The potential of the sphere is

$$\varphi = \frac{Q}{4\pi\epsilon_0 r}$$

and, therefore, the capacitance is

$$C = \frac{Q}{\varphi} = 4\pi\epsilon_0 r$$

9 Find the capacitance of a cylindrical capacitor of length l, when the radii of the plates are r_1 and r_2, and the space between the plates is filled with a dielectric of permittivity ϵ.

Solution: If the charge on the inside plate is Q, then using Gauss' theorem we find the electric field between the plates

$$E_r = \frac{Q}{2\pi\epsilon r l} \quad (r_1 < r < r_2)$$

From the potential difference between the plates, we find

$$\varphi_1 - \varphi_2 = \int_{r_1}^{r_2} E_r \, dr = \frac{1}{2\pi\epsilon} \frac{Q}{l} \ln \frac{r_2}{r_1}$$

The capacitance is equal to

$$C = \frac{Q}{\varphi_1 - \varphi_2} = \frac{2\pi\epsilon l}{\ln \frac{r_2}{r_1}}$$

10 Find the capacitance of a cylindrical capacitor of length l, with two layers of dielectric. The dimensions are given in Fig. 25.

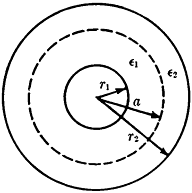


Fig. 25

Hint: Use the method of the preceding example.

$$\textit{Answer:} \quad C = \frac{2\pi l}{\frac{1}{\epsilon_1} \ln \frac{a}{r_1} + \frac{1}{\epsilon_2} \ln \frac{r_2}{a}}$$

11 Find the capacitance of a two-layer spherical capacitor. The dimensions are shown in Fig. 25.

$$\textit{Answer:} \quad C = \frac{4\pi}{\frac{1}{\epsilon_1}\left(\frac{1}{r_1} - \frac{1}{a}\right) + \frac{1}{\epsilon_2}\left(\frac{1}{a} - \frac{1}{r_2}\right)}$$

12 Determine the work done in moving apart the plates of a plane-parallel capacitor by a distance d. The area of each plate is S and the charge on it is Q.
Consider the numerical example: $Q = 10^{-10}$ coul, $S = 100$ cm^2, $d = 1$ cm.
Solution: The density of the force acting on a plate is given by equation (16.14).

The field on the plate surface can be calculated using the boundary condition (9.15)

$$E = \frac{\sigma}{\epsilon_0}$$

Consequently, the force is equal to

$$F = f_p S = \frac{1}{2}\epsilon_0 \frac{\sigma^2}{\epsilon_0^2} S = \frac{1}{2}\frac{Q^2}{\epsilon_0 S}$$

The work done in moving the plates apart is

$$A = Fd = \frac{1}{2}\frac{Q^2}{\epsilon_0 S} d \approx 5.6 \times 10^{-10}\ \mathrm{J}$$

13 The distance between the plates of a plane-parallel capacitor is d. Between the capacitor plates there is a metal plate of thickness δ, parallel to the capac-

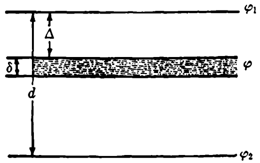


Fig. 26

itor plates. The capacitor plates are at potentials φ_1 and φ_2 (Fig. 26). Find the potential φ of the metal plate.

Solution: We have

$$\varphi_1 - \varphi = E\Delta \qquad \varphi - \varphi_2 = E(d - \Delta - \delta)$$

Hence

$$\varphi_1 - \varphi_2 = E(d - \delta)$$

Therefore

$$\varphi = \varphi_1 - E\Delta = \varphi_1 - \frac{\Delta}{d - \delta}(\varphi_1 - \varphi_2)$$

14 Determine the magnitude of the force acting on a charge q situated at a distance d from the center of an uncharged insulated conducting sphere of radius r_0 ($d > r_0$).

Answer: $F = -\dfrac{q^2 r_0^3}{4\pi\epsilon_0 d^3}\left[\dfrac{2d^2 - r_0^2}{(d^2 - r_0^2)^2}\right]$

15 Determine the force acting on a charge q placed inside a metal sphere at a distance r from the center. The radius of the sphere is a.

Answer: $F = \dfrac{q^2 a r}{4\pi\epsilon_0 (a^2 - r^2)^2}$

16 The distance between the plates of a plane-parallel capacitor is $d + a$ at one side, and $d - a$ at the other side, where $a \ll d$. Ignoring, as usual, the edge effects, determine, with accuracy to the order of $(a/d)^2$, the capacitance of the capacitor.

$$\textit{Answer:}\quad C = \frac{\epsilon_0 S}{d}\left(1 + \frac{1}{3}\frac{a^2}{d^2}\right)$$

where S is the area of the capacitor plates.

17 Consider two concentric conducting spheres with radii r_1 and r_2 $(r_1 < r_2)$. Between the spheres, at a distance d from their common center $(r_1 < d < r_2)$, there is a point charge q. Determine the charges induced on the spheres.

$$\textit{Answer:}\quad q_1 = -\frac{r_1(r_2 - d)}{d(r_2 - r_1)}\,q \qquad q_2 = -\frac{r_2(d - r_1)}{d(r_2 - r_1)}\,q$$

18 A small sphere is suspended above a horizontal conducting surface by an elastic thread which is an insulator. The coefficient of elasticity of the thread is equal to k. The distance from the sphere to the surface is d. Determine the charge q which must be given to the sphere to reduce by x the distance between the sphere and the conducting plane.

$$\textit{Answer:}\quad q = 4(d - x)\sqrt{kx\pi\epsilon_0}$$

19 It is found that on the surface of the earth the electric field is directed along the vertical and is approximately equal to 300 V/m. At a height of 1400 m above the earth, the vertical electric field equals about 20 V/m. Determine the surface charge density of the earth, considering the latter as a conductor. Determine the mean volume density of charge in the earth's atmosphere at heights below 1400 m.

$$\textit{Answer:}\quad \sigma = 2.65 \times 10^{-9}\ \text{coul/m}^2$$

$$\rho = 1.77 \times 10^{-12}\ \text{coul/m}^3$$

20 At a distance d from the center of an earthed sphere there is a point charge q. Determine the ratio f of the charge induced on the part of the sphere which subtends the point where the charge q is situated to the charge on the rest of the sphere. The radius of the sphere is a.

$$\textit{Answer:}\quad f = \sqrt{\frac{d + a}{d - a}}$$

21 Two infinite conducting plates are separated by a distance d. Between the plates, at a distance x from one of them, there is a point charge q. Determine the charges induced on the plates.

$$\textit{Answer:}\quad q_1 = -\frac{d - x}{d}\,q \qquad q_2 = -\frac{x}{d}\,q$$

22 Two capacitors of capacitances C_1 and C_2 and carrying charges q_1 and q_2 (the charge of a capacitor is the absolute magnitude of the charge on each plate of the capacitor) are connected in parallel. Calculate and explain the change in the electrostatic energy of the capacitors.

Answer: $\Delta W = \dfrac{(C_2q_1 - C_1q_2)^2}{2C_1C_2(C_1 + C_2)}$

23 In a medium of permittivity ϵ_1 there is an infinite straight filament, uniformly charged with a line density τ. This filament is parallel to the surface separating two dielectrics. The permittivity of the second dielectric is ϵ_2. The distance of the filament from the boundary surface is d. Determine the force per unit length of the filament.

Answer: $F_1 = \dfrac{\tau^2}{4\pi\epsilon_1}\left(\dfrac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}\right)\dfrac{1}{d}$

24 The permittivity of a medium between the plates of a plane-parallel capacitor varies uniformly as one moves from one plate of the capacitor to the other. The distance between the plates is d, and the values of the permittivity at the plates are ϵ_1 and ϵ_2. Find the capacitance of the capacitor.

Answer: $C = \dfrac{S}{d}\dfrac{\epsilon_2 - \epsilon_1}{\ln(\epsilon_2/\epsilon_1)}$

25 A cylindrical capacitor with radii of the plates r_1 and r_2 is immersed vertically in a liquid dielectric of permittivity ϵ. The lower end of the capacitor is in the liquid, and the upper end is in air, the permittivity of which is taken to be ϵ_0. The density of the liquid is ρ. Determine the height h to which the dielectric liquid rises between the plates of the capacitor, if the potential difference between the plates is V.

Answer: $h = \dfrac{(\epsilon - \epsilon_0)V^2}{(r_2^2 - r_1^2)\ln\dfrac{r_2}{r_1}}\dfrac{1}{\rho g}$

where g is the acceleration due to gravity.

26 A conducting sphere of density ρ_1 floats in a liquid of density ρ_2, the permittivity of which is equal to ϵ. Assume $\rho_2 > 2\rho_1$; the floating sphere is, therefore, less than half submerged in the liquid. What charge must be put on the sphere to make it half-submerged in the liquid? The radius of the sphere is r_0.

Answer: $q = 4\pi(\epsilon + \epsilon_0)\sqrt{\dfrac{r_0^5 g(\rho_2 - 2\rho_1)}{3(\epsilon - \epsilon_0)}}$

27 Consider a plane-parallel capacitor, the plates of which are squares of side a and a distance d apart. A dielectric plate of thickness Δ is moved into the space between the capacitor plates. This plate is also a square of side a. The surface of the dielectric plate is parallel to the surfaces of the capacitor plates, and the edges of the dielectric plate are parallel to the corresponding edges of the capacitor plates. Determine the force F with which the dielectric is attracted into the space between the condenser plates, if its permittivity is ϵ and the potential difference between the capacitor plates is V.

Answer: $F = \dfrac{\epsilon_0}{2}\dfrac{(\epsilon - \epsilon_0)\Delta}{(d - \Delta)\epsilon + \Delta\,\epsilon_0}\dfrac{a}{d}V^2$

CHAPTER 3

Static Magnetic Field

§17. General Properties and Equations of the Magnetostatic Field

Magnetostatics is the study of magnetic fields which are independent of time. Such fields are produced by constant currents. As was shown in § 10, in the case of electromagnetic fields which are time-independent, the magnetic field and the electric field may be considered independently of each other.

Mathematically, the domain of magnetostatics is characterized by the conditions: (1) all values are independent of time, (2) there are constant currents present. Hence, Maxwell's equations and the boundary conditions which apply to a magnetic field that is constant with respect to time have the form

$$\begin{aligned} \operatorname{curl} \mathbf{H} &= \mathbf{j} \qquad & B_{2n} - B_{1n} &= 0 \\ \operatorname{div} \mathbf{B} &= 0 \qquad & H_{2t} - H_{1t} &= i_{\text{surf}} \\ \mathbf{B} &= \mu \mathbf{H} \end{aligned} \tag{17.1}$$

The theory of the magnetic field must solve three fundamental problems: (1) given the field, to find the currents which produce it, (2) given the currents, to find the field which they produce, (3) to determine the forces acting in the magnetostatic field.

The first problem, as in the electrostatic case, is always solved easily, since the current for a given field can always be found by using Maxwell's equations and the boundary conditions (17.1)

$$\mathbf{j} = \operatorname{curl} \mathbf{H} \qquad i_{\text{surf}} = (H_{2t} - H_{1t})$$

The second and third problems are more interesting.

§18. Applied EMF's and the Generalized Ohm's and Joule-Lenz Laws

Impossibility of a Steady Current in the Presence of Coulomb Forces of the Electrostatic Field Only. In the case of steady currents, the equation of continuity (4.6) reduces to

$$\operatorname{div} \mathbf{j} = 0 \tag{18.1}$$

which expresses the fact that the lines of the current density vector have neither beginning nor end but are either closed lines or else go off to infinity.

We shall consider the integral

$$\oint_L \mathbf{j} \cdot d\mathbf{l} \tag{18.2}$$

taken along some closed contour L corresponding to some closed line of current. If there is only an electrostatic field $\mathbf{E}$ of coulomb origin, then the current density $\mathbf{j}$ is related to the field intensity $\mathbf{E}$ by the equation $\mathbf{j} = \lambda \mathbf{E}$, where $\mathbf{E} = -\operatorname{grad} \varphi$. Keeping this in mind, and taking $\lambda = \text{const}$, we obtain for (18.2)

$$\oint_L \mathbf{j} \cdot d\mathbf{l} = \lambda \oint_L \mathbf{E} \cdot d\mathbf{l} = -\lambda \oint_L \operatorname{grad} \varphi \, d\mathbf{l} = -\lambda \oint_L d\varphi = 0 \tag{18.3}$$

On the other hand, in (18.2), the integration is carried out along a current density line. Hence, an element of the path of integration $d\mathbf{l}$ is parallel to the current density vector at every point of the path. Consequently

$$\mathbf{j} \cdot d\mathbf{l} = \pm j \, dl$$

The plus sign is taken when $\mathbf{j}$ and $d\mathbf{l}$ are in the same direction, and the minus sign when they are in opposite directions. Then (18.3) becomes

$$\oint_L \mathbf{j} \cdot d\mathbf{l} = \pm \int_L j \, dl = 0$$

The integrand does not change sign along the whole path of integration. It therefore follows that since the integral is equal to zero, $\mathbf{j} = 0$. Thus, we have proved that it is impossible to have a steady current when only coulomb forces are present in the electrostatic field.

Applied Electromotive Forces. Steady currents can exist only in the presence of fields of noncoulomb origin. Such fields exist in steady current sources: galvanic cells, accumulators, etc. These are called *applied fields*, and the associated forces, which cause the motion of charge, are called *applied electromotive forces* (applied emf's).

Generalization of Ohm's Law. The current density is affected not only by the electrostatic field, but also by the field of the applied forces. The field of the applied forces is described by the vector $\mathbf{E}^{\text{app}}$, which is defined

as the electric field intensity, which would produce the same current density as that produced by the nonelectrostatic forces. Consequently, Ohm's law must be written in the form

$$\mathbf{j} = \lambda(\mathbf{E} + \mathbf{E}^{\mathrm{app}}) \tag{18.4}$$

Integral Form of Ohm's Law for a Complete Circuit. We shall multiply both sides of (18.4) by an element of length $d\mathbf{l}$ and integrate along a closed current density line in the direction of the current

$$\oint \mathbf{j}\cdot d\mathbf{l} = \lambda \oint (\mathbf{E} + \mathbf{E}^{\mathrm{app}})\cdot d\mathbf{l}$$

Dividing both sides of the equation by λ, and taking (18.3) into account, we obtain

$$\oint \frac{\mathbf{j}\cdot d\mathbf{l}}{\lambda} = \oint \mathbf{E}^{\mathrm{app}}\cdot d\mathbf{l} = \mathcal{E}^{\mathrm{app}} \tag{18.5}$$

Here, the notation

$$\mathcal{E}^{\mathrm{app}} = \oint \mathbf{E}^{\mathrm{app}}\cdot d\mathbf{l} \tag{18.6}$$

is introduced by analogy to the notation of the emf of electromagnetic origin

$$\mathcal{E}^{\mathrm{el}} = \oint \mathbf{E}\cdot d\mathbf{l}$$

Since $\mathbf{j}$ and $d\mathbf{l}$ in (18.5) lie along the same direction, we may write

$$\oint \frac{\mathbf{j}\cdot d\mathbf{l}}{\lambda} = \oint \frac{j\,dl}{\lambda} = \oint jS\frac{dl}{\lambda S} \tag{18.7}$$

where the numerator and denominator in the last integral have been multiplied by S, which is the cross section of a sufficiently small tube of current along which the integral (18.7) is taken. It is evident that the resistance of the section dl of the tube of current with the cross section S and electrical conductivity λ is equal to

$$dR = \frac{dl}{\lambda S} \tag{18.8}$$

while the current I flowing through the tube is equal to

$$I = jS$$

since in any cross section of the same tube of current, the current will be the same. Taking this into account, we may write

$$\oint jS\frac{dl}{\lambda S} = \int I\,dR = I\int dR = IR \tag{18.9}$$

Here, R is the resistance of the tube of current under consideration, and I is the current. On the basis of (18.9) and (18.7), equation (18.5) may be written in the form

$$IR = \mathcal{E}^{\text{app}} \tag{18.10}$$

This relationship is called Ohm's law. Equation (18.10) states that the existence and magnitudes of steady currents are, indeed, affected by the presence of applied emf's.

Generalization of the Joule-Lenz Law. The heat liberated by currents flowing in some volume V is given by the equation (3.7)

$$Q = \int_V \frac{j^2}{\lambda} dV$$

We transform the integrand in this formula using the generalized Ohm's law (18.4)

$$\frac{j^2}{\lambda} = \mathbf{j}\cdot\mathbf{E} + \mathbf{j}\cdot\mathbf{E}^{\text{app}}$$

As a result, we obtain

$$Q = \int_V \mathbf{j}\cdot\mathbf{E}\, dV + \int_V \mathbf{j}\cdot\mathbf{E}^{\text{app}}\, dV \tag{18.11}$$

It is easy to show that the first of these integrals is equal to zero. To prove this, we recall that $\mathbf{E} = -\text{grad}\, \varphi$, and use a well-known vector analysis formula (A.13) (in Appendix 1), which in this case is written

$$\mathbf{j}\cdot\mathbf{E} = -\mathbf{j}\,\text{grad}\, \varphi = -\text{div}\,(\varphi\mathbf{j}) + \varphi\, \text{div}\, \mathbf{j}$$

As a result, we obtain

$$\int_V \mathbf{j}\cdot\mathbf{E}\, dV = -\int_V \text{div}\,(\varphi\mathbf{j})\, dV + \int_V \varphi\, \text{div}\, \mathbf{j}\, dV \tag{18.12}$$

According to (18.1) the second integral is equal to zero, while the first integral may be transformed into a surface integral with the aid of Gauss' theorem

$$\int_V \text{div}\,(\varphi\mathbf{j})\, dV = \int_S \varphi\mathbf{j}\cdot d\mathbf{S}$$

We take all currents to be concentrated inside V. Then there are no currents flowing through the surface S which encloses V. This means that the current density on S is zero, and hence, the integral (18.12) equals zero. Equation (18.11) may, therefore, be written

$$Q = \int_V \mathbf{j}\cdot\mathbf{E}^{\text{app}}\, dV \tag{18.13}$$

This means that a steady current liberates heat solely due to the energy supplied to it by the applied emf's. The energy of the magnetic field remains constant.

§19. Magnetostatic Field in a Homogeneous Medium. Biot-Savart Law

Vector Potential. Maxwell's equation

$$\text{curl}\,\mathbf{H} = \mathbf{j} \tag{19.1}$$

states that in contrast to the case of an electrostatic field, a magnetostatic field is not, generally speaking, a potential field, while the equation

$$\text{div}\,\mathbf{B} = 0 \tag{19.2}$$

indicates that there are no magnetic charges to create a magnetostatic field in the way electric charges create an electrostatic field.

It is known from vector analysis that the divergence of the curl of any vector is identically equal to zero. Hence, it follows that a general solution of (19.2) may be obtained by writing **B** as the curl of some vector **A**

$$\mathbf{B} = \text{curl}\,\mathbf{A} \tag{19.3}$$

A is called the *vector potential of the magnetic field,* or simply the *vector potential.*

The vector potential is not a unique function of a given magnetic field **B**. If a given magnetic field **B** is described by a potential **A** according to (19.3), then the potential

$$\mathbf{A}' = \mathbf{A} + \text{grad}\,\chi \tag{19.4}$$

where χ is an arbitrary function, will also describe **B** since

$$\mathbf{B}' = \text{curl}\,\mathbf{A}' = \text{curl}\,(\mathbf{A} + \text{grad}\,\chi) = \text{curl}\,\mathbf{A} + \text{curl grad}\,\chi = \text{curl}\,\mathbf{A} = \mathbf{B} \tag{19.5}$$

Here, we take into account that the curl of the gradient always equals zero. This means that the potentials **A** and **A′**, which differ by the gradient of an arbitrary function, describe the same magnetic field.

Making use of this arbitrariness of the choice of the potential, we can impose some additional condition on the potential. In magnetostatics, we take as our arbitrary condition

$$\text{div}\,\mathbf{A} = 0 \tag{19.6}$$

The vector potential is an auxiliary function and has no physical meaning.

Vector Potential. The equation for the vector potential is obtained by substituting the expression (19.3) for $\mu\mathbf{H}$ in (19.1)

$$\text{curl curl}\,\mathbf{A} = \mu\mathbf{j} \tag{19.7}$$

We use the vector analysis formula (A.8) in Appendix 1

$$\text{curl curl}\,\mathbf{A} = \text{grad div}\,\mathbf{A} - \Delta\mathbf{A} \tag{19.8}$$

Since, according to (19.6), the first term on the right-hand side of (19.8)

is equal to zero, we obtain the following equation for the vector potential

$$\Delta \mathbf{A} = -\mu \mathbf{j} \tag{19.9}$$

This equation may also be written in the form of three scalar equations for the components of the vectors

$$\Delta A_x = -\mu j_x \qquad \Delta A_y = -\mu j_y \qquad \Delta A_z = -\mu j_z$$

Thus, the vector potential obeys Poisson's equation, and the solution may be written, by analogy to the solution of Poisson's equation for the scalar potential, in the form

$$A_x = \frac{\mu}{4\pi}\int_V \frac{j_x\,dV}{r} \qquad A_y = \frac{\mu}{4\pi}\int_V \frac{j_y\,dV}{r} \qquad A_z = \frac{\mu}{4\pi}\int_V \frac{j_z\,dV}{r}$$

or in the vector form

$$\mathbf{A} = \frac{\mu}{4\pi}\int_V \frac{\mathbf{j}\,dV}{r} \tag{19.10}$$

Here, $\mathbf{A}$ is the value of the vector potential at a point P where the field is measured, $\mathbf{j}$ is the current density in an element dV of the volume of integration, r is the distance between dV and P.

Biot-Savart Law. The magnitude of the magnetic field is evaluated according to (19.3) using the vector potential (19.10). Substituting (19.10) in (19.3), we obtain

$$\mathbf{B} = \operatorname{curl}\mathbf{A} = \frac{\mu}{4\pi}\operatorname{curl}\int_V \frac{\mathbf{j}\,dV}{r} = \frac{\mu}{4\pi}\int_V \operatorname{curl}\left(\frac{\mathbf{j}}{r}\right)dV \tag{19.11}$$

where the curl operator is taken into the integrand because the region of integration does not depend on the parameters upon which the curl operator acts, i.e., it is independent of the coordinates of the point where the vector potential is calculated.

To transform the integrand in (19.11), we use formula (A.14) in Appendix 1 written in the form

$$\operatorname{curl}\left(\frac{1}{r}\mathbf{j}\right) = \frac{1}{r}\operatorname{curl}\mathbf{j} + \operatorname{grad}\frac{1}{r}\times\mathbf{j} \tag{19.12}$$

In the integrand of (19.11), the current density vector $\mathbf{j}$ depends on the coordinates of the point of integration, while the curl operation is carried out with respect to the coordinates of the point at which the field is measured, i.e., with respect to completely different, independent, variables.

Consequently, $\operatorname{curl}\mathbf{j}$ in the first term of the right-hand side of (19.12) must be put equal to zero in the substitution in the integrand of (19.11). Remembering also that

$$\operatorname{grad}\frac{1}{r}\times\mathbf{j} = -\frac{\mathbf{r}}{r^3}\times\mathbf{j} = \frac{\mathbf{j}\times\mathbf{r}}{r^3} \tag{19.13}$$

and substituting (19.12) in (19.11), we obtain the Biot-Savart law

$$\mathbf{B} = \frac{\mu}{4\pi} \int_V \frac{\mathbf{j} \times \mathbf{r}}{r^3}\, dV \tag{19.14}$$

The radius vector in this equation is the radius vector taken from the element of the volume of integration to the point where the field is calculated.

Equation (19.14) solves the problem of finding the magnetic field produced by given currents. Taking into account the fact that $\mathbf{B} = \mu\mathbf{H}$, equation (19.14) may be put in the form

$$\mathbf{H} = \frac{1}{4\pi} \int_V \frac{\mathbf{j} \times \mathbf{r}}{r^3}\, dV \tag{19.15}$$

Hence, it is clear that the magnitude of the intensity vector **H** of the magnetic field produced by a given distribution of conduction currents is independent of the medium: a given conduction current will produce the same magnetic field in a medium and *in vacuo*. This means that, in magnetic phenomena, the vector **H** plays the same part as does the vector **D** in the electric field theory. From this point of view, it would be more correct to call **H** the magnetic induction vector, but this name belongs, historically, to **B**, although the latter plays a role in the magnetic field theory analogous to that of **E** in the electric field theory. It follows that the analogue of permittivity ϵ in the electric field theory is $1/\mu$ and not μ in the magnetic field theory.

Line Currents. In the majority of cases of practical importance, steady currents flow in thin conductors and are distributed with a uniform density over the cross sections of these conductors. Such currents are called *line currents*. A conductor is assumed to be thin if the linear dimensions of its cross section are much less than distances to the points at which the field is calculated.

The Biot-Savart law (19.14) simplifies in the case of line currents. Let us consider an element dl of the conductor. Its volume is equal to

$$dV = S\, dl$$

where S is the area of the cross section. Hence, we may write

$$\mathbf{j}\, dV = \mathbf{j}S\, dl = jS\, d\mathbf{l} = I\, d\mathbf{l} \tag{19.16}$$

where $I = jS$ is the current flowing along the conductor, and $d\mathbf{l}$ is an element of length of the conductor. The transformation (19.16) includes integration over the cross section of the conductor. Since we have assumed the dimensions of the cross section to be much less than distances to the points at which the field is measured, we can ignore, in the process of

integration, the differences between distances to the different elements of the cross section under consideration, and we may take this distance to be constant. Hence, transition to the case of line currents in (19.14) and (19.15) is effected by the simple change

$$\mathbf{j}\,dV \to I\,d\mathbf{l} \tag{19.16a}$$

in the integrand of the corresponding formula, taking r to be constant. We must now integrate with respect to $d\mathbf{l}$ along the whole conductor. In this integration, the current I is constant since the current is the same at all cross sections of the conductor, and hence, I may be taken outside the integral sign. Thus, (19.14) assumes the form

$$\mathbf{B} = \frac{\mu}{4\pi}\int_V \frac{\mathbf{j}\,dV \times \mathbf{r}}{r^3} = \frac{\mu}{4\pi} I \oint_L \frac{d\mathbf{l} \times \mathbf{r}}{r^3} \tag{19.17}$$

where V is the volume of the conductor and L is the linear contour of the conductor. Then, in the case of linear currents, the problem of determining the magnetic field is reduced, using the Biot-Savart formula, to that of evaluating a line integral around the contour of the conductor

$$\mathbf{H} = \frac{I}{4\pi}\oint \frac{d\mathbf{l} \times \mathbf{r}}{r^3} \tag{19.18}$$

Field of Elementary Currents. *Elementary currents* is the name given to closed currents flowing in a region of linear dimensions much less than the distance from that region to the points at which the field is calculated. Within certain limits, any steady current is elementary. This is the reason why a special study must be made of elementary currents.

The vector potential of an elementary current (Fig. 27) is evaluated by the equation

$$\mathbf{A} = \frac{\mu}{2\pi}\int \frac{\mathbf{j}\,dV}{r'} \tag{19.19}$$

The various quantities in the above formula are defined in Fig. 27. The origin is taken at the point O, and the field is calculated at the point Γ; V is the volume of the conductor carrying the current. The mathematical expression of the fact that the dimensions of the region carrying the current are small in comparison with the distance to the point Γ is given by

$$\frac{r_0}{r'} \ll 1 \qquad \frac{r_0}{r} \ll 1 \tag{19.20}$$

From Fig. 27, it is clear that

$$\mathbf{r}' = \mathbf{r} - \mathbf{r}_0$$

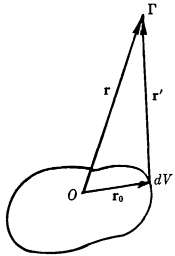


Fig. 27

Hence, it follows that

$$\frac{1}{r'} = \frac{1}{r}\left(1 - 2\frac{\mathbf{r}\cdot\mathbf{r}_0}{r^2} + \frac{r_0^2}{r^2}\right)^{-1/2} \tag{19.21}$$

Taking (19.20) into account, the right-hand side of (19.21) may be expanded in series, taking only the terms that are linear in r_0/r_1

$$\left(1 - 2\frac{\mathbf{r}\cdot\mathbf{r}_0}{r^2} + \frac{r_0^2}{r^2}\right)^{-1/2} \approx 1 + \frac{\mathbf{r}\cdot\mathbf{r}_0}{r^2}$$

Thus, we obtain the equation

$$\frac{1}{r'} = \frac{1}{r} + \frac{\mathbf{r}\cdot\mathbf{r}_0}{r^3} + \cdots$$

which, substituted in (19.19), leads to the following equation for the vector potential

$$\mathbf{A} = \frac{\mu}{4\pi}\frac{1}{r}\int_V \mathbf{j}\,dV + \frac{\mu}{4\pi}\frac{1}{r^3}\int_V \mathbf{j}\;(\mathbf{r}\cdot\mathbf{r}_0)\,dV \tag{19.22}$$

Here, we have allowed for the fact that $\mathbf{r}$ remains constant during the integration.

The first integral in the right-hand side of (19.22) is equal to zero. To verify this, it is simplest to divide up the region of integration into a number of separate current tubes. Applying (19.6) to the integration along each tube of current, we obtain

$$\int_V \mathbf{j}\,dV = I\oint d\mathbf{l} = 0$$

since each tube of current is closed, and, therefore,

$$\oint d\mathbf{l} = 0$$

To transform the second integral on the right-hand side of (19.22), we use the well-known formula of vector algebra, (A.1) in Appendix 1, for the expansion of a triple vector product

$$(\mathbf{r}_0 \times \mathbf{j}) \times \mathbf{r} = \mathbf{j}\ (\mathbf{r}\cdot\mathbf{r}_0) - \mathbf{r}_0\ (\mathbf{j}\cdot\mathbf{r})$$

This formula may be rewritten in the form

$$\mathbf{j}\ (\mathbf{r}\cdot\mathbf{r}_0) = \frac{1}{2}(\mathbf{r}_0 \times \mathbf{j}) \times \mathbf{r} + \frac{1}{2}\{\mathbf{j}\ (\mathbf{r}\cdot\mathbf{r}_0) + \mathbf{r}_0\ (\mathbf{j}\cdot\mathbf{r})\}$$

Substituting this expression in (19.22), we obtain

$$\mathbf{A} = \frac{\mu}{8\pi r^3}\int_V (\mathbf{r}_0 \times \mathbf{j}) \times \mathbf{r}\, dV + \frac{\mu}{8\pi r^3}\int_V \{\mathbf{j}\ (\mathbf{r}\cdot\mathbf{r}_0) + \mathbf{r}_0\ (\mathbf{j}\cdot\mathbf{r})\}\, dV \quad (19.23)$$

We shall show that the integral

$$\mathbf{K} = \int_V \{\mathbf{j}\ (\mathbf{r}\cdot\mathbf{r}_0) + \mathbf{r}_0\ (\mathbf{j}\cdot\mathbf{r})\}\, dV$$

is equal to zero. We multiply both sides by an arbitrary constant vector $\mathbf{a}$

$$\mathbf{a}\cdot\mathbf{K} = \int_V \{(\mathbf{a}\cdot\mathbf{j})\ (\mathbf{r}\cdot\mathbf{r}_0) + (\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{j}\cdot\mathbf{r})\}\, dV \quad (19.24)$$

The integrand may be transformed by means of the following equations

$$\begin{aligned}(\mathbf{a}\cdot\mathbf{j})\ (\mathbf{r}\cdot\mathbf{r}_0) + (\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{j}\cdot\mathbf{r}) &= \mathbf{j}\cdot\operatorname{grad}\{(\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{r}\cdot\mathbf{r}_0)\} \\ &= \operatorname{div}\{\mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{r}\cdot\mathbf{r}_0)\} - (\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{r}\cdot\mathbf{r}_0)\operatorname{div}\mathbf{j} \quad (19.25)\end{aligned}$$

where the first equation is verified by the formula for the gradient of a scalar product, and the second equation is the result of applying formula (A.13) in Appendix 1. Substituting (19.25) in (19.24), and remembering that $\operatorname{div}\mathbf{j} = 0$, we find

$$\mathbf{a}\cdot\mathbf{K} = \int_V \operatorname{div}\{\mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{r}\cdot\mathbf{r}_0)\}\, dV = \int_S d\mathbf{S}\cdot\mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0)\ (\mathbf{r}\cdot\mathbf{r}_0) \quad (19.26)$$

where we have used Gauss' theorem, and taken into account that all the currents are concentrated inside the given volume V and that, therefore, the current density at points on the surface bounding this volume is equal to zero, and that hence, the integral over this surface in (19.26) is equal to zero.

From the equation

$$\mathbf{a}\cdot\mathbf{K} = 0 \quad (19.27)$$

it follows that, since $\mathbf{a}$ is arbitrary

$$\mathbf{K} = 0 \quad (19.28)$$

The proof of the latter assertion is obvious if we consider the converse. If $\mathbf{K} \neq 0$, then we may choose $\mathbf{a}$ to be a vector parallel to $\mathbf{K}$. Then

$$\mathbf{a}\cdot\mathbf{K} = \pm aK \neq 0$$

which contradicts (19.27). Hence, the assumption $\mathbf{K} \neq 0$ leads to a contradiction and is, therefore, untrue.

Thus, taking (19.28) into account, we find

$$\mathbf{A} = \frac{\mu}{8\pi r^3} \int (\mathbf{r}_0 \times \mathbf{j}) \times \mathbf{r}\, dV = \left(\frac{\mu}{8\pi r^3} \int_V \mathbf{r}_0 \times \mathbf{j}\, dV\right) \times \mathbf{r} \quad (19.29)$$

taking note of the fact that $\mathbf{r}$ remains constant during the integration.

The vector

$$\mathbf{M} = \frac{1}{2} \int_V \mathbf{r}_0 \times \mathbf{j}\, dV \quad (19.30)$$

is called the *magnetic moment of an elementary current*. From (19.29) and (19.30), it is clear that the vector potential and, consequently, the magnetic field of a closed current loop at a sufficiently great distance from it, is completely defined by the magnetic moment of this current

$$\mathbf{A} = \frac{\mu}{4\pi} \frac{\mathbf{M} \times \mathbf{r}}{r^3} \quad (19.31)$$

$$\mathbf{B} = \frac{\mu}{4\pi} \operatorname{curl} \frac{\mathbf{M} \times \mathbf{r}}{r^3} = \frac{\mu}{4\pi} \left\{\frac{3(\mathbf{M}\cdot\mathbf{r})\,\mathbf{r}}{r^5} - \frac{\mathbf{M}}{r^3}\right\} \quad (19.32)$$

Equation (19.32) states that the magnetic field of a closed current loop decreases in inverse proportion to the cube of the distance.

Magnetic Moment of a Line Current. In the case of a line current, equation (19.30) for the magnetic moment takes a simpler form. Using the substitution (19.16a), we obtain

$$\mathbf{M} = \frac{1}{2} \int_V \mathbf{r}_0 \times \mathbf{j}\, dV = \frac{I}{2} \int_L \mathbf{r}_0 \times d\mathbf{l} \quad (19.33)$$

where L is the contour of the given closed current loop (Fig. 28). We use the fact that

$$\frac{1}{2} \mathbf{r}_0 \times d\mathbf{l} = d\mathbf{S} \quad (19.34)$$

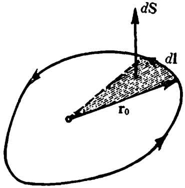


Fig. 28

is the vector of an element of the surface bounded by the current contour, and hence

$$\frac{1}{2}\int_L \mathbf{r}_0 \times d\mathbf{l} = \int_S d\mathbf{S} = \mathbf{S} \tag{19.35}$$

is the vector area of the surface subtended by the current contour L. Substituting (19.35) in (19.33) we find

$$\mathbf{M} = I\mathbf{S} \tag{19.36}$$

The surface vector $\mathbf{S}$ makes a right-hand screw system with the direction of the flow of current round the contour.

§20. Magnetic Substances in a Magnetic Field

Magnetic substances are substances which are capable of affecting a magnetic field, either producing it or changing it.

Magnetization of Magnetic Substances. When a magnetic substance is placed in a magnetic field, it acquires a magnetic moment or, as we say, it becomes magnetized. The intensity of magnetization is described by a vector $\mathbf{I}$, which is defined as the magnetic moment per unit volume of the magnetic substance. Thus, the magnetic moment $d\mathbf{M}$ of an element of volume dV of a magnetic substance, whose intensity of magnetization is $\mathbf{I}$, equals

$$d\mathbf{M} = \mathbf{I}\, dV \tag{20.1}$$

The presence of a magnetic moment in every volume element creates, in accordance with (19.32), an additional magnetic field, which is superimposed on the external field. Thus, the effect of magnetization of magnetic materials on the magnetic field is analogous to that of polarization of dielectrics on the electric field. However, there is a very important difference. In a dielectric, the additional field is always directed opposite to the original external field, and, hence, the total field in the dielectric is always less than the original field. In magnetic substances, the additional field may have either the same direction or the opposite direction as the initial field, according to the properties of the magnetic substance. Magnetic substances in which the additional field is in the opposite direction to the original field are called *diamagnetics*. Magnetic substances in which the additional field has the same direction as the original field are called *paramagnetics*. Thus, diamagnetics weaken the original field, and paramagnetics strengthen it. For all diamagnetics, and most paramagnetics, the additional magnetic field is very small in comparison with the original external field. When the original field disappears, the additional field also

disappears, i.e., diamagnetics and paramagnetics become completely demagnetized. But, there is also a third class of magnetic substances, whose additional field is much greater than the original field, and this additional field does not disappear when the original external field disappears. Consequently, these magnetic substances exhibit *residual magnetism.* They are able not only to change a magnetic field, but to create one by themselves. Such substances are called *ferromagnetics.* It is impossible to construct a strict theory of the magnetization of ferromagnetic substances within the framework of classical electrodynamics, since their magnetization is caused by the spin magnetism of the electrons, in the description of which a large part is played by the quantum laws. Hence, the theory of magnetic substances discussed in this course on electrodynamics is applicable only to diamagnetics and paramagnetics. Some remarks on ferromagnetics will be given in a special section in the second part of the book.

The magnitude of the magnetization vector $\mathbf{I}$ is related to the original external magnetic field by the equation

$$\mathbf{I} = \chi \mathbf{H} \tag{20.2}$$

The coefficient χ is called the *coefficient of magnetic susceptibility*, and may be either positive or negative.

Vector Potential in the Presence of Magnetic Substances. The effect of a magnetic substance on a magnetic field is to create an additional field due to the magnetization of the magnetic substance. Therefore, the total magnetic field in the presence of a magnetic substance is the sum of two fields: (1) the magnetic field of conduction currents (the vector potential of which is denoted by $\mathbf{A}_0$), (2) the magnetic field set up by the magnetization of the magnetic substance (the vector potential of which is $\mathbf{A}_M$).

Hence, the vector potential $\mathbf{A}$ of the total magnetic field is given by

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_M \tag{20.3}$$

where

$$\mathbf{A}_0 = \frac{\mu_0}{4\pi} \int_V \frac{\mathbf{j}\, dV}{r} \tag{20.4}$$

Here $\mathbf{j}$ is the density of the conduction current flowing in V.

From equation (19.31) it follows that the vector potential $d\mathbf{A}_M$ set up by the magnetic moment $d\mathbf{M}$ is given by

$$d\mathbf{A}_M = \frac{\mu_0}{4\pi} \frac{d\mathbf{M} \times \mathbf{r}}{r^3} \tag{20.5}$$

Hence, taking (20.1) into account, we obtain

$$d\mathbf{A}_M = \frac{\mu_0}{4\pi} \frac{\mathbf{I} \times \mathbf{r}}{r^3}\, dV \tag{20.6}$$

and, consequently

$$\mathbf{A}_M = \frac{\mu_0}{4\pi}\int_V \frac{\mathbf{I}\times\mathbf{r}}{r^3}\,dV \tag{20.7}$$

Thus, the problem of determining the field due to the magnetization of the magnetic substance is solved.

For convenience, we may put (20.7) into another form. For this purpose we use a well-known formula of vector analysis (A.14) of Appendix 1, which in this case has the form

$$\operatorname{curl}\left(\frac{\mathbf{I}}{r}\right) = \frac{1}{r}\operatorname{curl}\mathbf{I} + \operatorname{grad}\frac{1}{r}\times\mathbf{I} \tag{20.8}$$

We apply this formula to transform the integrand in (20.7), taking into account that the radius vector $\mathbf{r}$ in (20.7) is directed from the volume element dV to the point where the vector potential $\mathbf{A}_M$ is evaluated. It is clear that in equation (20.8), all the operations of differentiation must be taken as operations with respect to the same variables. Thus, taking the operations curl and gradient as operations with respect to the coordinates of the volume element dV, we have

$$\operatorname{curl}\left(\frac{\mathbf{I}}{r}\right) = \frac{1}{r}\operatorname{curl}\mathbf{I} + \frac{\mathbf{r}\times\mathbf{I}}{r^3} = \frac{1}{r}\operatorname{curl}\mathbf{I} - \frac{\mathbf{I}\times\mathbf{r}}{r^3} \tag{20.9}$$

Here $\mathbf{r}$ is the radius vector from the volume element dV to the point where the field is evaluated.

Using (20.9), equation (20.7) for the potential may be transformed into

$$\mathbf{A}_M = \frac{\mu_0}{4\pi}\int\frac{\operatorname{curl}\mathbf{I}}{r}\,dV - \frac{\mu_0}{4\pi}\int\operatorname{curl}\left(\frac{\mathbf{I}}{r}\right)dV \tag{20.10}$$

In further transformation of the second integral in (20.10), we use the following vector analysis formula

$$\int_V \operatorname{curl}\mathbf{A}\,dV = \oint_S d\mathbf{S}\times\mathbf{A} \tag{20.11}$$

where $\mathbf{A}$ is a vector function continuous in V, and the vector $d\mathbf{S}$ is taken along the outward normal to the surface S bounding V.

Equation (20.2) indicates that the magnetization vector $\mathbf{I}$ has discontinuities at the boundaries between different magnetic substances and at the boundary between a magnetic substance and empty space. Hence, in order to apply equation (20.11) to transform the second integral in the right-hand side of (20.10) we must exclude the boundary of discontinuity of the vector function $\mathbf{I}$, as we have already done several times (Fig. 16). Now let the surface S'' in Fig. 16 be the surface enclosing the volume under consideration, and let S be the surface separating magnetic substances, on

which $\mathbf{I}$ is discontinuous. We exclude this surface S from the region of integration by means of a surface S'. The volume V under consideration is now bounded by the surfaces S'' and S'. In this volume $\mathbf{I}$ is continuous, and we may, therefore, apply equation (20.10)

$$\int_V \operatorname{curl}\left(\frac{\mathbf{I}}{r}\right) dV = \int_{S''} \frac{d\mathbf{S} \times \mathbf{I}}{r} + \int_{S'} \frac{d\mathbf{S} \times \mathbf{I}}{r} \tag{20.12}$$

If we assume that the positive normal to S points in the direction of substance 2, as shown in Fig. 16, and shrink S' to S, we obtain

$$\int_{S'} \frac{d\mathbf{S} \times \mathbf{I}}{r} = -\int_S \frac{d\mathbf{S} \times \mathbf{I}_2}{r} + \int_S \frac{d\mathbf{S} \times \mathbf{I}_1}{r}$$

$$= \int_S \frac{d\mathbf{S} \times (\mathbf{I}_1 - \mathbf{I}_2)}{r} = \int_S \frac{\mathbf{n} \times (\mathbf{I}_1 - \mathbf{I}_2)}{r} dS \tag{20.13}$$

where $\mathbf{n}$ is a unit vector normal to S, pointing towards substance 2. If we assume that all the magnetic substances lie within V, so that S'' does not intersect them, then $\mathbf{I} = 0$ in the integrand of the first integral of the right-hand side of (20.12). Hence, the first integral in the right-hand side of (20.12) vanishes. The second integral is transformed using equation (20.13). Thus, taking (20.13) and (20.12) into account, equation (20.10) for the vector potential may be put in the following form

$$\mathbf{A}_M = \frac{\mu_0}{4\pi} \int_V \frac{\operatorname{curl} \mathbf{I}}{r} dV + \frac{\mu_0}{4\pi} \int_S \frac{\mathbf{n} \times (\mathbf{I}_2 - \mathbf{I}_1)}{r} dS \tag{20.14}$$

where S is taken to mean the sum of all the surfaces separating magnetic materials on which $\mathbf{I}$ is discontinuous.

It follows from Maxwell's equations that there are no magnetic charges. A magnetic field can be created only by currents. Hence, the magnetization which leads to the creation of an additional magnetic field must be associated with the creation of some currents. However, in contrast to conduction currents, associated with the displacement of charges through macroscopic distances, these currents are associated with the motion of charges through microscopic distances, i.e., they are associated with the motion of charges in molecules. They are, therefore, called *molecular currents*. Thus, the magnetization is caused by molecular currents. We must point out that this applies only to paramagnetics and diamagnetics. The magnetic properties of ferromagnetics are due to the magnetic properties of electrons, and cannot be explained by molecular currents.

We compare the first term of the right-hand side of equation (20.14) with the expression for the vector potential (20.4). We see that in equation (20.14) for the vector potential arising as a result of the magnetization,

curl **I** plays the part of the volume current density. Consequently, the mean volume density of the molecular currents $\langle \mathbf{j}_{\text{mol}} \rangle$ is given by

$$\langle \mathbf{j}_{\text{mol}} \rangle = \text{curl}\, \mathbf{I} \tag{20.15}$$

Here, we speak of the mean volume density, since the true volume density of the molecular currents, obviously, changes very rapidly as we pass from atom to atom, and the intensity of magnetization **I** represents some mean magnetization, caused by these molecular currents.

Thus, the first term in (20.14) represents the creation of a magnetic field due to the presence of a mean volume density of molecular currents. Therefore, it follows that the second term in (20.14) describes the creation of a magnetic field due to the presence of a mean surface density of molecular currents $\langle \mathbf{i}_{\text{surf mol}} \rangle$. Hence

$$\langle \mathbf{i}_{\text{surf mol}} \rangle = \mathbf{n} \times (\mathbf{I}_2 - \mathbf{I}_1) \tag{20.16}$$

and equation (20.14) for the vector potential due to the magnetization of a magnetic substance may be put in the following form

$$\mathbf{A}_M = \frac{\mu_0}{4\pi} \int_V \frac{\langle \mathbf{j}_{\text{mol}} \rangle}{r}\, dV + \frac{\mu_0}{4\pi} \int_S \frac{\langle \mathbf{i}_{\text{surf mol}} \rangle}{r}\, dS \tag{20.17}$$

On the basis of equation (20.17), we may say that the magnetic field created by a magnetic substance is produced by volume and surface molecular currents of the magnetic substance.

The Relationship Between Permeability and Magnetic Susceptibility. As has been shown, the presence of magnetic substances may be completely described, if, in addition to the magnetic field produced by conduction currents, we also take into account the magnetic field produced by molecular currents.

The field created by conduction currents *in vacuo* is described by Maxwell's equation

$$\text{curl}\, \mathbf{B} = \mu_0 \mathbf{j} \tag{20.18}$$

In order to allow for the presence of magnetic substances, it is necessary to include in (20.18) not only the conduction currents **j**, but also the molecular currents of (20.15). Hence, if **B** is taken to be the magnetic induction in the presence of magnetic substances, equation (20.18) must be written

$$\text{curl}\, \mathbf{B} = \mu_0(\mathbf{j} + \text{curl}\, \mathbf{I}) \tag{20.19}$$

Taking curl **I** over to the left-hand side of (20.19), and dividing both sides by μ_0, we obtain

$$\text{curl}\left(\frac{\mathbf{B}}{\mu_0} - \mathbf{I}\right) = \mathbf{j} \tag{20.20}$$

On the other hand, Maxwell's equation (17.1)

$$\operatorname{curl} \mathbf{H} = \mathbf{j} \tag{20.21}$$

holds, whether or not magnetic substances are present. But equations (20.21) and (20.20) describe the same magnetic field, and are different forms of the same equation. Comparison of these equations gives

$$\frac{\mathbf{B}}{\mu_0} - \mathbf{I} = \mathbf{H} \tag{20.22}$$

This last equation may be considered as the definition of the vector **H**, which is called the magnetic field in a medium.

Substituting

$$\mathbf{B} = \mu \mathbf{H} \qquad \mathbf{I} = \chi \mathbf{H} \tag{20.23}$$

in (20.22), we obtain the following relationship between the permeability μ and the magnetic susceptibility

$$\mu = \mu_0(1 + \chi) \qquad \chi = \frac{\mu - \mu_0}{\mu_0} \tag{20.24}$$

The magnetic susceptibility in the Gaussian system of units is expressed by a quantity 4π times smaller

$$\chi' = \frac{\chi}{4\pi}$$

Hence, substituting this expression for χ', and the relationship

$$\mu = \mu' \mu_0 \tag{20.25}$$

in (20.24), we obtain the following relationship between the permeability and the magnetic susceptibility in the Gaussian system of units

$$\mu' = 1 + 4\pi\chi' \tag{20.26}$$

The value of χ may be either positive or negative, and, consequently, the permeability of a magnetic substance may be either greater or less than the permeability of empty space. For diamagnetics

$$\chi < 0 \qquad \mu < \mu_0 \tag{20.27}$$

For paramagnetics

$$\chi > 0 \qquad \mu > \mu_0 \tag{20.28}$$

In ferromagnetics, the permeability depends on the field and, moreover, these substances possess *residual magnetism*; in fact, in ferromagnetics $\mu \gg \mu_0$. The theory given in this section does not apply to ferromagnetics.

Magnetic Field of Permanent Magnets. As we have pointed out, the explanation of the nature of ferromagnetism lies outside the framework

of classical electrodynamics. The magnetic field of ferromagnetics (e.g., permanent magnets) is caused by the spin magnetism of the electrons, and cannot be interpreted as being due to the existence of currents in the electron. If we assume this interpretation, then to obtain intelligible quantitative results, we must take the linear velocity of rotation of the electrons to be greater than the velocity of light, which contradicts the theory of relativity. Hence, we must reject the interpretation of the spin magnetism of the electron as being caused by circular currents within the electron. However, the magnetic field due to the spin magnetism may be described approximately within the framework of phenomenological electrodynamics.

The intensity of magnetization of a magnetic substance is described by the magnetization vector $\mathbf{I}$. In the case of permanent magnets, the magnetization is also described by the magnetization vector $\mathbf{I}_0$, without discussing the nature of this magnetization and the causes of permanent magnetization. The permanent magnetization $\mathbf{I}_0$ produces the same magnetic field as an equal induced magnetization would. Hence, we may use equation (20.14) to determine the magnetic field of permanent magnets, and this gives the following expression for the vector potential $\mathbf{A}_p$ of a permanent magnet

$$\mathbf{A}_p = \frac{\mu_0}{4\pi}\int_V \frac{\operatorname{curl}\mathbf{I}_0}{r}\,dV + \frac{\mu_0}{4\pi}\int_S \frac{\mathbf{I}_0 \times \mathbf{n}}{r}\,dS \tag{20.29}$$

The permanent magnet magnetization differs from the induced magnetization only in the fact that, in the case of permanent magnets, the value of $\mathbf{I}_0$ is not governed by an external field, according to (20.2), but is independent of the external field. Using (20.14), the magnetic field of permanent magnets is brought into the scheme of steady magnetic fields discussed in this chapter. To make the analogy more complete, we may assume the vector $\mathbf{I}_0$ of a permanent magnet to be due to effective molecular currents, remembering that these are fictitious currents. The equation for the magnetic induction in this case is still

$$\operatorname{div}\mathbf{B} = 0 \tag{20.30}$$

but that Maxwell's equation, which expresses the relationship between the magnetic induction and the magnetic field, is somewhat altered. Since a permanent magnet is an additional source of the magnetic field, we may write, instead of (20.22), the equation

$$\mathbf{B} = \mu_0\mathbf{H} + \mu_0\mathbf{I} + \mu_0\mathbf{I}_0 \tag{20.31}$$

Since

$$\mu_0\mathbf{H} + \mu_0\mathbf{I} = \mu\mathbf{H} \tag{20.32}$$

we finally obtain

$$\mathbf{B} = \mu\mathbf{H} + \mu_0\mathbf{I}_0 \tag{20.33}$$

The above equations describe completely the field of permanent magnets.

§21. Energy of the Magnetic Field of Steady Currents

Expression for the Energy of a Magnetic Field in Terms of Field Vectors. From the general expression for the energy of an electromagnetic field (8.12), it follows that

$$W = \frac{1}{2}\int_V \mathbf{H}\cdot\mathbf{B}\,dV \tag{21.1}$$

This equation states that the energy of a magnetic field is distributed in space with the density

$$U_{\mathrm{M}} = \frac{1}{2}\mathbf{H}\cdot\mathbf{B} \tag{21.2}$$

Expression for the Energy in Terms of the Current Density and the Vector Potential. Using the expression $\mathbf{B} = \operatorname{curl}\mathbf{A}$, together with the well-known vector analysis formula (A.15) of Appendix 1, the integrand of (21.1) may be rearranged

$$\mathbf{H}\cdot\mathbf{B} = \mathbf{H}\cdot\operatorname{curl}\mathbf{A} = \operatorname{div}\mathbf{A}\times\mathbf{H} + \mathbf{A}\cdot\operatorname{curl}\mathbf{H} = \operatorname{div}\mathbf{A}\times\mathbf{H} + \mathbf{A}\cdot\mathbf{j} \tag{21.3}$$

where Maxwell's equation $\operatorname{curl}\mathbf{H} = \mathbf{j}$ is used in the last equation. Consequently, for the energy of a magnetic field, we obtain, instead of (21.1), the expression

$$W = \frac{1}{2}\int_V \operatorname{div}\mathbf{A}\times\mathbf{H}\,dV + \frac{1}{2}\int_V \mathbf{A}\cdot\mathbf{j}\,dV \tag{21.4}$$

The first integral on the right-hand side of this expression is transformed into a surface integral using Gauss' theorem

$$\int_V \operatorname{div}\mathbf{A}\times\mathbf{H}\,dV = \oint_S \mathbf{A}\times\mathbf{H}\cdot d\mathbf{S} \tag{21.5}$$

If all currents are in a finite region of space, then at great distances from them [see equation (19.31)] we have

$$A \sim \frac{1}{r^2} \qquad H \sim \frac{1}{r^3}$$

At large distances, therefore, the integrand decreases as does $1/r^5$. Since the surface of integration increases only as does r^2, the integral decreases, when $S \to \infty$, as does $1/r^3$. Hence, as the surface S moves away to infinity,

the integral (21.5) tends to zero. Therefore, equation (21.4) takes the form

$$W = \frac{1}{2}\int_V \mathbf{A}\cdot\mathbf{j}\, dV \tag{21.6}$$

This formula expresses the energy of a magnetic field as the energy of interaction between a current and a field described by $\mathbf{A}$.

Expression for the Energy of a Magnetic Field as the Energy of Interaction Between Current Elements. If in (21.6), the potential $\mathbf{A}$ is expressed in terms of the currents

$$\mathbf{A} = \frac{\mu}{4\pi}\int_V \frac{\mathbf{j}'\, dV'}{r} \tag{21.7}$$

we obtain

$$W = \frac{\mu}{8\pi}\int_V \frac{\mathbf{j}\cdot\mathbf{j}'}{r}\, dV\, dV' \tag{21.8}$$

where r is the distance between the volume elements dV and dV'. This formula expresses the energy of a magnetic field as the energy of interaction between current elements.

Energy of a Magnetic Field for a System of Line Currents. We now proceed to the case of line currents using rule (19.16a). We take equation (21.6) as the initial expression. The integral in (21.6) reduces to a sum of integrals over the volumes of separate conductors, designated by the index k. We thus obtain

$$W = \frac{1}{2}\int_V \mathbf{A}\cdot\mathbf{j}\, dV = \frac{1}{2}\sum_k \int_{V_k} \mathbf{A}\cdot\mathbf{j}\, dV$$
$$= \frac{1}{2}\sum_k \int_{L_k} \mathbf{A}I_k\cdot d\mathbf{l} = \frac{1}{2}\sum_k I_k \int_{L_k} \mathbf{A}\cdot d\mathbf{l} \tag{21.9}$$

Here, V_k and L_k are, respectively, the volume and contour of the k^{th} conductor, and I_k is the total current flowing in the k^{th} conductor. The summation is carried out over all conductors. Using Stokes' theorem

$$\oint_{L_k} \mathbf{A}\cdot d\mathbf{l} = \int_{S_k} \operatorname{curl}\mathbf{A}\cdot d\mathbf{S} = \int_{S_k} \mathbf{B}\cdot d\mathbf{S} = \Phi_k \tag{21.10}$$

The quantity Φ_k is the magnetic induction flux through a surface S_k subtended by the contour L_k of the k^{th} conductor. Thus, the expression for the energy of the magnetic field of a system of line currents may be written

$$W = \frac{1}{2}\sum_k I_k\Phi_k \tag{21.11}$$

In particular, in the presence of a single conductor, this equation takes the form

$$W = \frac{1}{2} I\Phi \tag{21.12}$$

Expression for the Coefficients of Self-Induction and Mutual Induction. The integration in equation (21.8) reduces to the integration over the volumes of the conductors. Then, denoting the volume of the k^{th} conductor by V_k, we have

$$W = \frac{\mu}{8\pi} \int_V \frac{\mathbf{j} \cdot \mathbf{j}'}{r} \, dV \, dV' = \frac{\mu}{8\pi} \sum_{i,k} \int_{V_i} \int_{V_k} \frac{\mathbf{j}_i \cdot \mathbf{j}_k}{r_{ik}} \, dV_i \, dV_k \tag{21.13}$$

where $\mathbf{j}_i$, $\mathbf{j}_k$ denote the current densities in the i^{th} and j^{th} conductors, and r_{ik} is the distance between the volume elements dV_i and dV_k of the two conductors. Equation (21.13) may be rewritten in the form

$$W = \frac{\mu}{8\pi} \sum_{i,k} I_k I_i \frac{1}{I_i I_k} \int_{V_i} \int_{V_k} \frac{\mathbf{j}_i \cdot \mathbf{j}_k}{r_{ik}} \, dV_i \, dV_k = \frac{1}{2} \sum_{i,k} L_{ik} I_i I_k \tag{21.14}$$

where

$$L_{ik} = \frac{\mu}{4\pi} \frac{1}{I_i I_k} \int_{V_i} \int_{V_k} \frac{\mathbf{j}_i \cdot \mathbf{j}_k}{r_{ik}} \, dV_i \, dV_k \tag{21.15}$$

The coefficients L_{ik} depend only on the shape of conductors and their relative positions, but are independent of the currents flowing in the conductors because, when the current changes, the numerator and denominator in equation (21.15) change by the same factor. For a given system of conductors, therefore, the coefficients L_{ik} may be evaluated once for all. If we know these coefficients, and the currents flowing in conductors, we can use (21.14) to calculate the magnetic field energy of a system of conductors and currents. The coefficient L_{ik} for $i \neq k$ is called the *coefficient of mutual induction* of the i^{th} and k^{th} conductors, and for $i = j$ it is known as the *coefficient of self-induction* of a given conductor.

Equation (21.14) may be used to determine the coefficient of self-induction of a conductor. For a single insulated conductor, this expression takes the form

$$W = \frac{1}{2} LI^2 \tag{21.16}$$

If we can measure or calculate the energy W, then L may be found from the above equation. This method is often used.

In the calculation of the coefficients of mutual induction L_{ik} in the case

of line currents, one may use the transformation (19.16a) in (21.15). We thus obtain

$$L_{ik} = \frac{\mu}{4\pi} \int_{L_i} \int_{L_k} \frac{d\mathbf{l}_i \cdot d\mathbf{l}_k}{r_{ik}} \qquad i \neq k \tag{21.17}$$

where L_i and L_k are the contours of the i^{th} and k^{th} currents. This is called *Neumann's formula.* This expression cannot be used to calculate the coefficient of self-induction, since the integral of (21.17) would become infinite. It follows from (21.15) and (21.17) that

$$L_{ik} = L_{ki} \tag{21.18}$$

Relationship Between the Coefficients of Self-Induction and Mutual Induction and Magnetic Induction Flux. Comparison of equation (21.11) with (21.14) written in the form

$$W = \frac{1}{2} \sum_k I_k \sum_i L_{ki} I_i$$

shows that

$$\Phi_k = \sum_i L_{ki} I_i \tag{21.19}$$

Thus, the knowledge of the coefficients of mutual induction and self-induction allows one to calculate the magnetic induction flux very easily.

Equation (21.19) is especially suitable in the case of a system of conductors for which the coefficients of mutual induction and self-induction are known and tabulated. For a single insulated conductor (21.19) takes the form

$$\Phi = LI \tag{21.20}$$

Hence, it follows that

$$L = \frac{\Phi}{I} \tag{21.21}$$

Thus, if we know the magnetic induction flux set up by a known current flowing through a conductor, we can use (21.21) to calculate the coefficient of self-induction of the conductor. Equation (21.21) is often used to calculate the coefficients of self-induction. The inductance is measured in henrys (H).

Energy of a Magnetic Moment. The expression for the energy of a magnetic moment in a magnetic field may be found conveniently if we start from the forces acting on the magnetic moment. This expression will be deduced in the following section in a discussion of mechanical forces.

§22. Mechanical Forces in the Magnetostatic Field

Force on an Element of Current. *Ampère's law* expresses the fact that the force $d\mathbf{F}$ on an element of current $\mathbf{j}\,dV$ in a field $\mathbf{B}$ is equal to

$$d\mathbf{F} = \mathbf{j} \times \mathbf{B}\,dV \tag{22.1}$$

In the case of line currents, this expression takes the form

$$d\mathbf{F} = I\,d\mathbf{l} \times \mathbf{B} \tag{22.2}$$

where $d\mathbf{l}$ is an element of length of the conductor.

To find the force acting on a finite segment of the conductor, we must integrate

$$\mathbf{F} = \int_V \mathbf{j} \times \mathbf{B}\,dV \tag{22.3}$$

$$\mathbf{F} = I \int_L d\mathbf{l} \times \mathbf{B} \tag{22.4}$$

where V and L are, respectively, the volume and contour of the conductor or a part of it.

Force on an Elementary Current. In our discussion of the magnetic field set up by an elementary current, we have found that, in the first approximation, the field was completely described by the magnetic moment of the current. We shall now show that the force on an elementary current in a given magnetic field is also completely described, in the first approximation, by the magnetic moment of this current. We must now add an additional condition to the definition of an elementary current given above: the linear dimensions of the region in which the elementary current flows are so small that the change in the external magnetic field in this region is small and we need only the first-order terms with respect to the linear dimensions of the region. Let us take some point close to the elementary current as the origin of coordinates. By the definition of an elementary current, near the origin, we may expand $\mathbf{B}$ as a Taylor series, taking only the linear terms

$$\mathbf{B}(x, y, z) = \mathbf{B}(0, 0, 0) + x\frac{\partial \mathbf{B}}{\partial x} + y\frac{\partial \mathbf{B}}{\partial y} + z\frac{\partial \mathbf{B}}{\partial z} + \cdots \tag{22.5}$$

The derivatives in this expansion are taken at the point $x = y = z = 0$. Taking (22.5) into account, (22.3) becomes

$$\mathbf{F} = \int_V \mathbf{j} \times \mathbf{B}_0\,dV + \int_V \mathbf{j} \times \left(x\frac{\partial \mathbf{B}}{\partial x} + y\frac{\partial \mathbf{B}}{\partial y} + z\frac{\partial \mathbf{B}}{\partial z}\right) dV \tag{22.6}$$

where $\mathbf{B}_0 = \mathbf{B}\,(0, 0, 0)$. The first integral is equal to zero, since $\mathbf{B}\,(0, 0, 0)$ is constant, and, therefore

$$\int_V \mathbf{j} \times \mathbf{B}_0 \, dV = \left(\int_V \mathbf{j} \, dV\right) \times \mathbf{B}_0 = 0 \tag{22.7}$$

because, for a closed current loop

$$\int_V \mathbf{j} \, dV = 0 \tag{22.8}$$

We shall consider one of the components of $\mathbf{F}$, e.g., the z component

$$\begin{aligned} F_z &= \int_V \left[\mathbf{j} \times \left(x \frac{\partial \mathbf{B}}{\partial x} + y \frac{\partial \mathbf{B}}{\partial y} + z \frac{\partial \mathbf{B}}{\partial z}\right)\right]_z dV \\ &= \int_V \left\{j_x \left(x \frac{\partial B_y}{\partial x} + y \frac{\partial B_y}{\partial y} + z \frac{\partial B_y}{\partial z}\right) - j_y \left(x \frac{\partial B_x}{\partial x} + y \frac{\partial B_x}{\partial y} + z \frac{\partial B_x}{\partial z}\right)\right\} dV \end{aligned} \tag{22.9}$$

We shall show first of all that

$$\int_V xj_x \, dV = \int_V yj_y \, dV = \int_V zj_z \, dV = 0 \tag{22.10}$$

We shall consider the first integral as an example. We have

$$xj_x = \frac{1}{2} \mathbf{j} \cdot \operatorname{grad} x^2 = \frac{1}{2} \operatorname{div} (\mathbf{j}x^2) - \frac{1}{2} x^2 \operatorname{div} \mathbf{j} \tag{22.11}$$

where formula (A.13) of Appendix 1 is used. Since $\operatorname{div} \mathbf{j} = 0$, and using Gauss' theorem, we obtain

$$\int_V xj_x \, dV = \frac{1}{2} \int_V \operatorname{div} (\mathbf{j}x^2) \, dV = \frac{1}{2} \int_S \mathbf{j}x^2 \cdot d\mathbf{S} = 0 \tag{22.12}$$

because all the currents are concentrated in the volume V, and $\mathbf{j} = 0$ on the surface S which encloses V. The other equalities of (22.10) are proved in a similar manner.

We can now show that

$$\left.\begin{aligned} \int_V xj_y \, dV &= -\int_V yj_x \, dV = M_z \\ \int_V yj_z \, dV &= -\int_V zj_y \, dV = M_x \\ \int_V zj_x \, dV &= -\int_V xj_z \, dV = M_y \end{aligned}\right\} \tag{22.13}$$

where M_x, M_y, M_z are the components of the magnetic moment vector of the elementary current, defined by (19.30).

We consider the first equation of (22.13) as an example. We may write

$$xj_y = \frac{1}{2} (xj_y - yj_x) + \frac{1}{2} (xj_y + yj_x) \tag{22.14}$$

But

$$\frac{1}{2}(xj_y + yj_x) = \frac{1}{2}\mathbf{j}\cdot\mathrm{grad}\,(xy) = \frac{1}{2}\mathrm{div}\,(\mathbf{j}xy) - \frac{1}{2}xy\,\mathrm{div}\,\mathbf{j} \quad (22.15)$$

where we have used formula (A.13) of Appendix 1. Since div $\mathbf{j} = 0$, we obtain

$$\int_V xj_y\,dV = \frac{1}{2}\int_V (xj_y - yj_x)\,dV + \frac{1}{2}\int_V \mathrm{div}\,(\mathbf{j}xy)\,dV$$

$$= \frac{1}{2}\int_V (xj_y - yj_x)\,dV + \frac{1}{2}\int_S \mathbf{j}xy\cdot d\mathbf{S} \quad (22.16)$$

The last integral is equal to zero by the considerations used in proving (22.12). Hence, we finally obtain

$$\int_V xj_y\,dV = \frac{1}{2}\int_V (xj_y - yj_x)\,dV = M_z \quad (22.17)$$

The other equations are proved in a similar manner.

Using (22.10) and (22.13), equation (22.9) may be rewritten as

$$F_z = -M_z\frac{\partial B_y}{\partial y} + M_y\frac{\partial B_y}{\partial z} - M_z\frac{\partial B_x}{\partial x} + M_x\frac{\partial B_x}{\partial z} \quad (22.18)$$

Maxwell's equation

$$\mathrm{div}\,\mathbf{B} = \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0 \quad (22.19)$$

shows that

$$-M_z\frac{\partial B_y}{\partial y} - M_z\frac{\partial B_x}{\partial x} = M_z\frac{\partial B_z}{\partial z} \quad (22.20)$$

and, therefore, equation (22.18) takes the form

$$F_z = M_x\frac{\partial B_x}{\partial z} + M_y\frac{\partial B_y}{\partial z} + M_z\frac{\partial B_z}{\partial z} = \mathbf{M}\cdot\frac{\partial \mathbf{B}}{\partial z} \quad (22.21)$$

Formulas for the other two components may be obtained in a similar manner. Hence, we may write

$$F_x = \mathbf{M}\cdot\frac{\partial \mathbf{B}}{\partial x} \qquad F_y = \mathbf{M}\cdot\frac{\partial \mathbf{B}}{\partial y} \qquad F_z = \mathbf{M}\cdot\frac{\partial \mathbf{B}}{\partial z} \quad (22.22)$$

Thus, the force on an elementary current in a given magnetic field is completely defined by the magnetic moment of this current. The force increases with the inhomogeneity of the magnetic field. In a homogeneous magnetic field, this force is equal to zero.

Force on a Magnetic Substance. When placed in a magnetic field, a

magnetic substance becomes magnetized. An element of volume dV of the magnetic substance acquires a magnetic moment

$$d\mathbf{M} = \mathbf{I}\, dV \tag{22.23}$$

where $\mathbf{I}$ is the magnetization vector. According to equation (22.22), the force on this volume element of the magnetic substance is equal to

$$dF_x = \mathbf{I} \cdot \frac{\partial \mathbf{B}}{\partial x}\, dV \qquad dF_y = \mathbf{I} \cdot \frac{\partial \mathbf{B}}{\partial y}\, dV \qquad dF_z = \mathbf{I} \cdot \frac{\partial \mathbf{B}}{\partial z}\, dV \tag{22.24}$$

Taking into account that, on the basis of (20.24)

$$\mathbf{I} = \frac{\mu - \mu_0}{\mu\mu_0}\mathbf{B} \tag{22.25}$$

and substituting this value for $\mathbf{I}$ in equation (22.24), we find

$$\left.\begin{aligned} dF_x &= \frac{\mu - \mu_0}{\mu\mu_0}\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial x}\, dV = \frac{1}{2}\frac{\mu - \mu_0}{\mu\mu_0}\frac{\partial B^2}{\partial x}\, dV \\ dF_y &= \frac{\mu - \mu_0}{\mu\mu_0}\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial y}\, dV = \frac{1}{2}\frac{\mu - \mu_0}{\mu\mu_0}\frac{\partial B^2}{\partial y}\, dV \\ dF_z &= \frac{\mu - \mu_0}{\mu\mu_0}\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial z}\, dV = \frac{1}{2}\frac{\mu - \mu_0}{\mu\mu_0}\frac{\partial B^2}{\partial z}\, dV \end{aligned}\right\} \tag{22.26}$$

These three equations may be written in the form of a single vector equation

$$d\mathbf{F} = \frac{1}{2}\frac{\mu - \mu_0}{\mu\mu_0}\,\mathrm{grad}\, B^2\, dV \tag{22.27}$$

Thus, the volume density of the force $\mathbf{f}$ on a magnetic substance is equal to

$$\mathbf{f} = \frac{d\mathbf{F}}{dV} = \frac{1}{2}\frac{\mu - \mu_0}{\mu\mu_0}\,\mathrm{grad}\, B^2 \tag{22.28}$$

Comparison of this expression with equation (16.20) for the density of the force on a dielectric confirms the correctness of the postulate that the roles of $\mathbf{E}$ and ϵ in the theory of the electric field are taken, in magnetic field theory, by $\mathbf{B}$ and $1/\mu$, and not by $\mathbf{H}$ and μ.

Equation (22.28) shows that paramagnetics and diamagnetics behave differently in a magnetic field: the forces act on them in opposite directions. Since for paramagnetics $\mu > \mu_0$, the volume density of the force $\mathbf{f}$ is in the same direction as the gradient of the square of the magnetic induction. This means that paramagnetics are attracted towards regions of maximum induction. For diamagnetics $\mu < \mu_0$, and, therefore, diamagnetics are repelled from regions of stronger induction to regions of weaker induction. For example, a copper rod, being a paramagnetic, is attracted into a solenoid in which a current is flowing, but a bismuth rod, being a diamagnetic, is repelled from the solenoid.

Ferromagnetics, for which $\mu \gg \mu_0$, are also attracted into regions of

stronger induction, but the force is much stronger than for paramagnetics, because μ of paramagnetics is only slightly different from μ_0, and, therefore, the term $(\mu - \mu_0)/\mu$ in equation (22.28) is small. The strong force acting on ferromagnetics is used in various measuring devices.

Energy of a Magnetic Moment in an External Field. We shall consider the relationship between the forces acting on a system and the energy of a system (16.26). Since the magnetic moment **M** is not explicitly dependent on the coordinates, we obtain, from equations (16.26) and (22.22), the following expression for the energy of a magnetic moment in an external magnetic field

$$W = -\mathbf{M}\cdot\mathbf{B} \tag{22.29}$$

Moment of the Forces on a Magnetic Moment. If θ denotes the angle between **M** and **B**, then

$$W = -MB\cos\theta \tag{22.30}$$

As proved in theoretical mechanics, the generalized force corresponding to the generalized coordinate θ is a moment of forces (torque). Hence, the torque acting on a magnetic moment is equal to

$$N = -\frac{\partial W}{\partial \theta} = MB\sin\theta \tag{22.31}$$

or, in vector form

$$\mathbf{N} = \mathbf{M} \times \mathbf{B} \tag{22.32}$$

The torque tends to turn the magnetic moment **M** to coincide with **B**.

PROBLEMS

1 Determine the magnetic field on the axis of a circular current I of radius r_0 (Fig. 29).

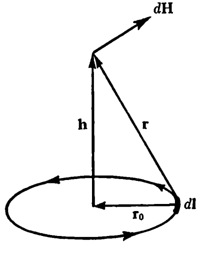


Fig. 29

Consider the numerical example: $I = 0.1$ amp, $\mathbf{r}_0 = 5$ cm, $h = 3$ cm.
Solution: We use the Biot-Savart law (19.18)

$$\mathbf{H} = \frac{I}{4\pi} \oint_L \frac{d\mathbf{l} \times \mathbf{r}}{r^3}$$

where $\mathbf{r} = \mathbf{r}_0 + \mathbf{h}$, $d\mathbf{l} \times \mathbf{r} = [d\mathbf{l} \times \mathbf{r}_0] + [d\mathbf{l} \times h]$. In the integration, the absolute magnitude of $\mathbf{r}$ does not change. Hence

$$H = \frac{I}{4\pi r^3}\left\{\oint_L d\mathbf{l} \times \mathbf{r}_0 + \oint_L d\mathbf{l} \times \mathbf{h}\right\}$$

Since $\mathbf{h}$ is a constant vector

$$\oint_L d\mathbf{l} \times \mathbf{h} = \left(\oint_L d\mathbf{l}\right) \times \mathbf{h} = 0$$

The second integral is evaluated as follows

$$\oint_L d\mathbf{l} \times \mathbf{r}_0 = \oint_L \mathbf{n} r_0\, dl = \mathbf{n} r_0 \oint_L dl = \mathbf{n} r_0 2\pi r_0$$

where $\mathbf{n}$ is a unit vector, parallel to $d\mathbf{l} \times \mathbf{r}_0$. Finally, we obtain

$$H_h = \frac{I}{2} \frac{r_0^2}{(r_0^2 + h^2)^{3/2}} = 6.2 \text{ amp/m}$$

The vector H is perpendicular to the plane of the ring.

2 Determine the magnetic field set up by a finite rectilinear current segment of length l (Fig. 30).

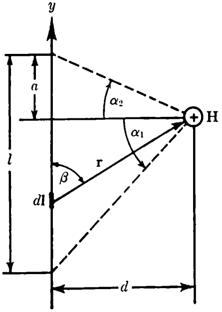


Fig. 30

Solution: The field $d\mathbf{H}$ due to each element of the conductor is perpendicular to the plane of Fig. 30

$$d\mathbf{H} = \frac{I}{4\pi}\frac{d\mathbf{l} \times \mathbf{r}}{r^3}$$

because the vector product $d\mathbf{l} \times \mathbf{r}$ is perpendicular to the plane of the figure. We have

$$|d\mathbf{l} \times \mathbf{r}| = dl\, r \sin (d\mathbf{l}, \mathbf{r}) = dl\, r \sin \beta = dy\, d$$

Hence

$$H = \frac{Id}{4\pi}\int_{-(l-a)}^{a} \frac{dy}{(d^2 + y^2)^{3/2}} = \frac{I}{4\pi d}(\sin \alpha_1 + \sin \alpha_2)$$

We may use this expression to calculate the field of any contour carrying a current, if the contour is made up of rectilinear segments.

3 Using the law of total current, find the magnetic field in a coaxial cable used for the transmission of a steady current I (Fig. 31). The current flows along

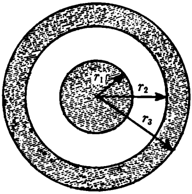


Fig. 31

the central conductor, radius r_1, and is enclosed by a shell of internal and external radii r_2 and r_3, respectively. The space between the central conductor and the shell is filled with a dielectric.

Solution: By axial symmetry and the law of total current, we have

$$H = \frac{i_r}{2\pi r}$$

where i_r is the current enclosed by a circular contour of radius r. The current density in the central conductor is equal to

$$j_1 = \frac{I}{\pi r_1^2}$$

Hence, for $0 < r < r_1$

$$i_r = \frac{I}{\pi r_1^2}\pi r^2 = I\frac{r^2}{r_1^2}$$

and, therefore

$$H = \frac{I}{2\pi}\frac{r}{r_1^2}$$

For $r_1 < r < r_2$, we have $i_r = I = \text{const}$, and, therefore

$$H = \frac{I}{2\pi}\frac{1}{r}$$

For $r_2 < r < r_3$, the contour encloses the return current, the density of which is

$$j_2 = \frac{I}{\pi(r_3^2 - r_2^2)}$$

Therefore, the total current enclosed by the contour $r_2 < r < r_3$ is written

$$i_r = I - I\frac{r^2 - r_2^2}{r_3^2 - r_2^2}$$

and the magnetic field is equal to

$$H = \frac{I}{2\pi r}\left(1 - \frac{r^2 - r_2^2}{r_3^2 - r_2^2}\right)$$

Outside the cable, the field is zero.

4 Calculate the inductance of a segment of length l of the coaxial cable of the previous problem, ignoring the internal inductance of the central conductor and the shell. The internal inductance of the central conductor is due to the interaction between different elements of current flowing along the central conductor. The internal inductance of the shell arises in a similar manner.
Solution: We use equation (21.21). The magnetic flux between the central conductor and the shell is calculated using the expression given in the preceding problem

$$\Phi = \mu_0\frac{I}{2\pi}l\int_{r_1}^{r_2}\frac{dr}{r} = \mu_0\frac{I}{2\pi}l\ln\frac{r_2}{r_1}$$

Hence, the inductance is

$$L = \frac{\Phi}{I} = \frac{\mu_0}{2\pi}l\ln\frac{r_2}{r_1}$$

5 Calculate the inductance of a segment of length l of a two-conductor line, ignoring the internal inductance of the conductors. The radii of the conductors are both equal to r_0; the distance between the conductors is d.
Hint: Use the method of the previous problem

Answer: $L = \frac{\mu_0}{\pi}l\ln\frac{d}{r_0}$

6 A current, of density $\mathbf{j}$, flows along an infinite circular cylindrical conductor. There is a circular cylindrical cavity in the conductor, and the axes of the cylinder and of the cavity are parallel (Fig. 24). Find the magnetic field inside the cavity ($\mu = \mu_0$).
Hint: Use an approach similar to that of Problem 6 of Chapter 2.

In the terminology of Fig. 24, we have

$$\mathbf{H} = \frac{1}{2}\mathbf{j} \times \mathbf{r}_0$$

The field outside the cavity is calculated in a similar manner.

7 A semicircular grounding electrode is buried in the ground, its top level with the ground surface (Fig. 32). Find the voltage V_s experienced by a man

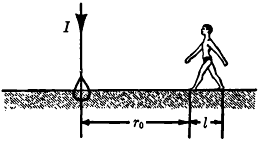


Fig. 32

walking up to the grounding electrode ("voltage per step"). The current I flowing through the grounding electrode is given. The length of the man's step is equal to l, and the distance of the man's nearer foot from the grounding electrode is r_0.

Consider the numerical example: $\lambda = 10^{-2}\ \text{ohm}^{-1}\ \text{m}^{-1}$, $I = 1$ amp, $r_0 = 2$ m, $l = 1$ m.

Solution: The current from the grounding electrode flows out uniformly in all directions, hence, the current density in the ground is equal to

$$j = \frac{I}{2\pi r^2}$$

From Ohm's law, we find the intensity of the electric field to be

$$E_r = \frac{j}{\lambda} = \frac{I}{2\pi r^2 \lambda}$$

Hence, the "voltage per step" is equal to

$$V_{\text{step}} = \int_{r_0}^{r_0+l} E_r\, dr = \frac{I}{2\pi\lambda}\int_{r_0}^{r_0+l}\frac{dr}{r^2} = \frac{I}{2\pi\lambda}\left(\frac{1}{r_0} - \frac{1}{r_0+l}\right) \approx 2.7\ \text{V}$$

8 An overhead conductor carrying a current I broke and landed on the ground. The length of the conductor on the ground is L. Find the "voltage per step" experienced by a man walking up to the conductor at right angles to the conductor. The length of the man's step is l. The distance of the foot nearer to the conductor is r_0. The conductivity of the earth $\lambda \approx 10^{-2}\ \text{ohm}^{-1}\ \text{m}^{-1}$.

Consider the numerical example: $r_0 = 1$ m, $l = 65$ cm, $L = 200$ m, $I = 500$ amp.

Solution: The solution is similar to that of the preceding problem

$$Answer: \quad V_{\text{step}} = \frac{I}{\pi\lambda L} \ln \frac{r_0 + l}{r_0} \approx 26.6 \text{ V}$$

9 A ring, of radius r_0, of very thin wire carries a current I. The mechanical strength of the wire is f_0. The ring is placed in a magnetic field perpendicular to the plane of the ring in such a way that the forces acting tend to break the ring. Find the magnitude of the magnetic field which will break the ring.
Consider the numerical example: $f_0 = 1.5$ N, $r_0 = 15$ cm, $I = 10$ amp.
Solution: The magnetic field force acts along the radius. Denoting an element of length of the ring by di, and using equation (22.2), we find the expression

$$dF = IB\,dl$$

for the absolute magnitude of an element of the force acting on the conductor. This force acts along the radius of the ring. Its projection on the x axis in the plane of the ring is

$$dF_x = dF \cos\alpha = IB\,dl \cos\alpha$$

where α is the angle between the x axis and the radius drawn to the element dl. Since $dl = r_0\,d\alpha$, the expression for the force acting on the half-ring in the x positive half-ring is

$$F_x = IBr_0 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\alpha\,d\alpha = 2IBr_0$$

This force is distributed over two cross sections of the conductor where it cuts the y axis. Hence, the condition for breaking the ring takes the form

$$2IBr_0 = 2f_0$$

and, therefore

$$B = \frac{f_0}{Ir_0} = 1 \text{ T} = 10^4 \text{ G}$$

10 A copper conductor of cross section S, bent in the form of three sides of a square of side l, may rotate about a horizontal axis passing through the fourth side of the square. The conductor is in a homogeneous vertical magnetic field. When a current I flows in the conductor, the latter turns through an angle α from the position of equilibrium. Given the value of the current and the angle of turn, determine the magnetic field.
Consider the numerical example: $S = 3$ mm², $I = 5$ amp, $\alpha = 15°$.
Solution: The forces on the horizontal part are equal to

$$F = IBl \qquad P = mg = \rho_m Slg$$

where ρ_m is the density of copper.
The condition of equilibrium is

$$\frac{F}{2P} = \tan\alpha$$

whence

$$B = \frac{2\rho_m S g}{I} \tan\alpha = 28\cdot 10^{-3}\,\mathrm{T} = 280\,\mathrm{G}$$

11 Find the field inside an infinite tightly wound solenoid (Fig. 33).
Solution: It is clear that, inside the solenoid, the magnetic field is directed along the solenoid axis, and that there is no radial component. We apply the law of total current to the dashed contour in Fig. 33. By the law of total

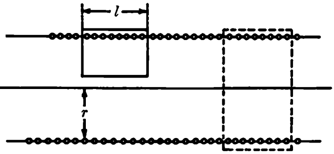


Fig. 33

current, the integral around this contour is equal to zero. Since the contribution of the vertical parts of the path of integration add up to zero, the integral reduces to the integrals along the parts outside the coil, parallel to the z axis. Along these parts, the directions of motion during integration are opposite. If there were a field component $H_z \neq 0$ acting, then it also would have opposite directions. Hence, the contributions of these two parts to the integral have the same sign. Since the integral is equal to zero, we must conclude that outside the solenoid $H_z = 0$. We now apply the law of total current to the (solid) contour of Fig. 33, and obtain

$$H_z l = NIl$$

where H_z is the magnetic field inside the solenoid, and N is the number of turns per unit length of the solenoid. Hence

$$H_z = NI$$

This equation is used in practical work for measuring of the magnetic field intensity in ampere-turns.

12 Determine the magnetic induction at the center of a square contour of side l carrying current I.

Answer: $B = \frac{2\sqrt{2}}{\pi}\mu_0 \frac{I}{l}\,\mathrm{T}$

13 A sphere of radius a is uniformly charged with a surface density σ, and rotates about its axis with an angular velocity ω. Find the magnetic induction at points inside the rotating sphere.

Answer: $B = \frac{2}{3}\mu_0\sigma a\omega$

14 A conductor is wound in the form of a helix on a cylindrical insulator, radius a, and makes n complete turns. The angle of the helix is α. Determine the magnetic induction at the center of the cylindrical insulator, when a current I flows in the coil.

$$\textit{Answer:}\quad B = \frac{\mu_0}{2}\frac{In}{a}\frac{1}{\sqrt{1+\pi^2n^2\tan^2\alpha}}$$

15 Consider a copper helix of radius a with n turns per meter. The weight of the helix may be neglected. The turns are wound so that there are very small gaps between them. The upper end of the helix is fixed, and the lower end is attached to a conducting load of mass m lying on a metal table. Assuming that the helix contracts uniformly, determine the current I which must flow in the helix in order to lift the weight off the table.

$$\textit{Answer:}\quad I = \frac{1}{na}\sqrt{\frac{2mg}{\pi\epsilon_0}}$$

16 Given a very long solenoid wound with n turns per unit length, and with a cross sectional area equal to S. A current I flows in the coil of the solenoid. Two very long iron rods of permeability μ are inserted into the solenoid, one from each end. The rods are inserted next to the solenoid coil, and there is a very small gap between the rods in the solenoid. Determine the force with which the rods attract each other.

$$\textit{Answer:}\quad F = \frac{S}{2}\frac{(\mu-\mu_0)\mu}{\mu_0}n^2I^2$$

17 Consider a U-shaped electromagnet, with a winding consisting of n turns. The cross sectional area, the length, the permeability of the material of the magnet, and the distance between the poles are S, l, μ, and d, respectively. The current flowing in the coil of the magnet is I. A "keeper" of the same material as the magnet, and with the same cross section, is placed across the poles. Find, approximately, the force with which the "keeper" is attracted to the magnet.

$$\textit{Answer:}\quad F = \frac{S}{(l+d)^2}\frac{\mu^2}{\mu_0}n^2I^2$$

18 Two small magnets of the same magnetic moment $|\mathbf{M}|$ and mass m are suspended on light threads. The distance between the points of suspension is very large. The lengths of the threads are equal. Show that the magnets will tend to be attracted to each other. Determine the angle of inclination θ of the threads to the vertical. The effect of the earth's magnetic field may be ignored.

$$\textit{Answer:}\quad \theta = \frac{3}{2}\frac{\mu_0}{\pi}\frac{M^2}{d^4}\frac{1}{mg}$$

CHAPTER 4

Quasi-Static Electromagnetic Fields

§23. Definitions and Equations

Definition. We say that an electromagnetic field is *quasi-static* if it changes very slowly with time. The criterion of "sufficient slowness" of the change in the field may be expressed as follows:

a) The change in the electromagnetic field is so slow that, in conducting media, the displacement current may be ignored in comparison with the conduction current

$$|\mathbf{j}_{\mathrm{disp}}|_{\mathrm{max}} \ll |\mathbf{j}|_{\mathrm{max}} \tag{23.1}$$

If the electromagnetic field varies with a frequency ω, i.e., if, for example

$$E = E_0 e^{i\omega t} \tag{23.2}$$

then

$$\begin{aligned} j_{\mathrm{disp}} &= \frac{\partial D}{\partial t} = i\omega\epsilon E_0 e^{i\omega t} \\ j &= \lambda E = \lambda E_0 e^{i\omega t} \end{aligned} \tag{23.3}$$

Hence, for (23.1) to be satisfied, the inequality

$$\frac{|j_{\mathrm{disp}}|_{\mathrm{max}}}{|j|_{\mathrm{max}}} = \frac{\omega\epsilon}{\lambda} \ll 1 \tag{23.4}$$

must hold. Since for metal conductors, $\epsilon \approx \epsilon_0$, $\lambda \approx 10^7$ ohm^{-1} cm^{-1}, we see that the displacement current is negligible in the range of frequencies

$$\omega \ll \frac{\lambda}{\epsilon_0} \approx 10^{18}\ \mathrm{sec}^{-1} \tag{23.5}$$

i.e., right up to ultraviolet frequencies. This estimate is approximate, since it does not allow for the inertial properties of a medium, which are

important at very high frequencies. If these inertial properties are taken into account, the estimate given by (23.5) is reduced by several orders of magnitude, but even then the range of frequencies in which the displacement current may be ignored in comparison with the conduction current is still very great.

b) The field changes so slowly that within the region of space under consideration we may ignore the lag effects due to the finite velocity of propagation of electromagnetic waves.

The change in the quantities which describe a plane wave propagated with a velocity c along the x axis may be written

$$E(x, t) = E_0 e^{i\omega\left(t - \frac{x}{c}\right)} = E_0 e^{i\omega t} e^{-i\omega \frac{x}{c}} \tag{23.6}$$

Expanding the last exponential factor as a series, we obtain

$$E(x, t) = E_0 e^{i\omega t}\left(1 - i\frac{\omega}{c}x + \cdots\right) \tag{23.7}$$

It is clear that the lag effects may be ignored whenever the dependence on x in the right-hand side of (23.7) may be ignored, i.e., when the inequality

$$\frac{\omega}{c}x \ll 1 \tag{23.8}$$

holds. Since

$$\frac{\omega}{c} = \frac{2\pi}{cT} = \frac{2\pi}{l} \tag{23.9}$$

where l is the wavelength, we may rewrite the condition (23.8) in the form

$$x \ll l \tag{23.10}$$

i.e., we may regard the velocity of propagation of electromagnetic waves to be infinite and we may ignore the lag effects if the linear dimensions of the region under consideration are many times smaller than the wavelength. If, for example, we deal with 50 cycle/sec current, then the corresponding wavelength is several thousand kilometers, and, therefore, the lag effects may be ignored, even over comparatively large regions.

From these considerations we conclude that the majority of fields used in power engineering and many of the fields used in radio engineering belong to the class of quasi-static fields.

Maxwell's Equations in a Quasi-Static Region. When displacement currents are ignored, Maxwell's equations take the form

$$\left.\begin{array}{ll} \operatorname{curl} \mathbf{H} = \mathbf{j} & \operatorname{curl} \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t} \\ \operatorname{div} \mathbf{B} = 0 & \operatorname{div} \mathbf{D} = \rho \end{array}\right\} \tag{23.11}$$

where

$$\mathbf{B} = \mu \mathbf{H} \qquad \mathbf{D} = \epsilon \mathbf{E} \qquad \mathbf{j} = \lambda(\mathbf{E} + \mathbf{E}^{\text{app}})$$

Thus, in the case of quasi-static fields, it is impossible to consider the electric and magnetic fields separately. However, only the principal relationship between them, due to Faraday's electromagnetic induction, is taken into account. The relationship involving displacement currents is of less importance, and it is ignored in the case of quasi-static fields.

Expression for the Electric Field Intensity in Terms of Potentials. Since the quasi-static electric field is due not only to the presence of charges, but also to changes in the magnetic field, the electric field depends not only on the scalar potential, but also on the vector potential. The vector potential **A** is introduced in the same way as in the case of a static magnetic field

$$\mathbf{B} = \operatorname{curl} \mathbf{A} \qquad \operatorname{div} \mathbf{A} = 0 \tag{23.12}$$

The electric field in the quasi-static case is not, generally speaking, a potential field, since

$$\operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \neq 0 \tag{23.13}$$

and, therefore, the electric intensity vector **E** cannot be written as the gradient of the scalar potential.

Expressing **B** in (23.13) in terms of the vector potential **A** we find

$$\operatorname{curl} \mathbf{E} = -\frac{\partial}{\partial t} \operatorname{curl} \mathbf{A} = -\operatorname{curl} \frac{\partial \mathbf{A}}{\partial t} \tag{23.14}$$

Here, the order of the operations of differentiation with respect to time and of taking the curl is reversed, because the curl operation consists of partial differentiation with respect to the coordinates: since the coordinates and time are independent variables, the order of differentiation may be reversed. Rewriting equation (23.14) in the form

$$\operatorname{curl}\left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0 \tag{23.15}$$

we see that the vector

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \tag{23.16}$$

is a potential vector, and, therefore, may be put in the form of the gradient of a scalar function

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\operatorname{grad} \varphi \tag{23.17}$$

Thus, the electric field vector is expressed in terms of scalar and vector potentials by the equation

$$\mathbf{E} = -\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} \tag{23.18}$$

The second term on the right-hand side of (23.18) takes into account Faraday's law of electromagnetic induction and is responsible for the nonpotential nature of the field in the quasi-static case. Due to the presence of this term, the work done by the field when a charge is moved from one point to another depends, generally speaking, on the form of the path.

Equation for the Scalar Potential. We shall consider the case of a homogeneous medium (ϵ = const). Substituting in

$$\operatorname{div} \mathbf{D} = \epsilon \operatorname{div} \mathbf{E} = \rho \tag{23.19}$$

the expression for $\mathbf{E}$ in terms of potentials, we find

$$\operatorname{div}\left(-\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t}\right) = \frac{\rho}{\epsilon} \tag{23.20}$$

Since $\operatorname{div} \operatorname{grad} \varphi = \nabla^2 \varphi$ and

$$\operatorname{div} \frac{\partial \mathbf{A}}{\partial t} = \frac{\partial}{\partial t} \operatorname{div} \mathbf{A} = 0 \tag{23.21}$$

we obtain the equation for the scalar potential in the form

$$\Delta \varphi = -\frac{\rho}{\epsilon} \tag{23.22}$$

This equation has exactly the same form as in the case of static fields. This is due to the fact that in the discussion of quasi-static fields we ignore the lag effects and assume that the scalar potential at a given point of space, at some instant of time, is determined by the distribution of charge throughout all space at that instant, and that the nature of the motion of the charge is of no importance. Therefore, the scalar potential has the same value which it would have if the charge were not moving.

Equation for the Vector Potential. All calculations and discussions are completely identical with those used in the deduction of the equation for the vector potential of a static magnetic field, since the initial equations are the same. Consequently, the equations for the vector potential are the same as before

$$\Delta \mathbf{A} = -\mu \mathbf{j} \tag{23.23}$$

In empty space, $\mu = \mu_0$ in this equation. The reason for the identity of this equation with the equation for a static magnetic field is the same as in the case of the scalar potential.

Thus, if we take a photograph of the instantaneous distribution of charges and currents at some instant of time, then we can determine the values of the scalar and vector potentials at all points of space. The values of the scalar and vector potentials will be the same as in the case of static fields which have the same distribution of charges and currents as that obtained in our instantaneous photograph. We can find the magnetic field from the value of the vector potential. We cannot obtain the value of the electric field, however, from a single instantaneous photograph, since the electric field depends on the derivative of the vector potential with respect to time. Hence, to determine the electric field we require a minimum of two instantaneous photographs of the distribution of currents, taken at infinitely close instants of time.

§24. System of Conductors, Taking Mutual Inductance and Self-Inductance into Account

The phenomenon of electromagnetic induction gives rise to an interaction between currents flowing in different conductors, and between elements of current flowing in the same conductor. Thus, we cannot consider a current flowing through some part of a circuit in isolation from the currents flowing in other parts of the circuit and in different circuits. We must consider all currents that are related to one another by induction.

Integral Ohm's Law Taking Electromagnetic Induction into Account. We shall consider a set of linear conductors. We shall apply Ohm's law in its differential form

$$\mathbf{j} = \lambda(\mathbf{E} + \mathbf{E}^{\text{app}}) \tag{24.1}$$

to the k^{th} conductor. We divide both sides of (24.1) by λ, multiply by the element of length $d\mathbf{l}$ of a linear conductor, and integrate around the closed contour of the conductor under consideration

$$\oint_{L_k} \frac{\mathbf{j}\cdot d\mathbf{l}}{\lambda} = \oint_{L_k} \mathbf{E}\cdot d\mathbf{l} + \oint_{L_k} \mathbf{E}^{\text{app}}\cdot d\mathbf{l} \tag{24.2}$$

Here, L_k is the contour of the k^{th} conductor. The integrand on the left-hand side may be transformed as follows

$$\frac{\mathbf{j}\cdot d\mathbf{l}}{\lambda} = \frac{j\,dl}{\lambda} = jS\frac{dl}{\lambda S} = I\,dR \tag{24.3}$$

where we take into account that in the conductor under consideration the direction of $\mathbf{j}$ may be taken to coincide with the direction of $d\mathbf{l}$; $dR = dl/\lambda S$ is the resistance of an element of length dl and cross section S; $I = jS$ is the current flowing in the conductor. Thus, we obtain

$$\int_{L_k} \frac{\mathbf{j}\cdot d\mathbf{l}}{\lambda} = \int_{L_k} I\, dR = I_k \int_{L_k} dR = I_k R_k \tag{24.4}$$

Here we allow for the fact that the current is the same across any cross section of a closed conductor, and, therefore, that I is constant during integration round the contour, and may, therefore, be taken outside the integral sign. R_k is the total resistance of the k^{th} conductor.

The integral

$$\int_{L_k} \mathbf{E}^{\text{app}}\cdot d\mathbf{l} = \mathcal{E}_k^{\text{app}} \tag{24.5}$$

on the right-hand side of (24.2) is the applied emf acting on the k^{th} conductor. The other integral on the right-hand side of (24.2) may be put in the following form

$$\oint_{L_k} \mathbf{E}\cdot d\mathbf{l} = -\oint_{L_k} \operatorname{grad} \varphi\cdot d\mathbf{l} - \oint \frac{\partial \mathbf{A}}{\partial t}\cdot d\mathbf{l} \tag{24.6}$$

using (23.18). Since L_k is closed, the first integral on the right-hand side of (24.6) equals zero

$$\oint_{L_k} \operatorname{grad} \varphi\cdot d\mathbf{l} = \oint_{L_k} d\varphi = 0$$

The second integral on the right-hand side of (24.6) may be transformed

$$\oint_{L_k} \frac{\partial \mathbf{A}}{\partial t}\cdot d\mathbf{l} = \frac{d}{dt}\oint_{L_k} \mathbf{A}\cdot d\mathbf{l} = \frac{d}{dt}\int_{S_k} \operatorname{curl} \mathbf{A}\cdot d\mathbf{S}$$

$$= \frac{d}{dt}\int_{S_k} \mathbf{B}\cdot d\mathbf{S} = \frac{d\Phi_k}{dt} \tag{24.7}$$

where

$$\Phi_k = \int_{S_k} \mathbf{B}\cdot d\mathbf{S}$$

is the magnetic induction flux through the surface S_k, subtended by the contour L_k of the k^{th} conductor.

In the transformation (24.7) we assume that the contour L_k is motionless, and, hence, the derivative with respect to time may be taken outside the integral sign. Stokes' theorem is also used.

Thus, using (24.4), (24.5) and (24.7), equation (24.2) may be rewritten

$$I_k R_k = \mathcal{E}_k^{\text{app}} - \frac{d\Phi_k}{dt} \tag{24.8}$$

This is Ohm's law for the k^{th} conductor, allowing for the electromagnetic induction which is described by the second term in the right-hand side of this equation.

Equation for a System of Conductors. The magnetic induction flux Φ_k through the surface subtended by the k^{th} conductor may be written, according to (21.19), in the form

$$\Phi_k = \sum_{i=1}^{N} L_{ki} I_i \tag{24.9}$$

where the summation with respect to i is carried out over all conductors (assuming that there are N of them).

Substituting this expression for Φ_k in (24.8), we obtain a system of ordinary differential equations

$$I_k R_k = \mathcal{E}_k^{\text{app}} - \sum_{i=1}^{N} L_{ki} \frac{dI_i}{dt} \qquad k = 1, 2, \ldots N \tag{24.10}$$

Here, we take the coefficients L_{ki} to be constants, and replace partial derivatives with total derivatives, since in these equations the only variable is time, and the currents I_i depend on it explicitly. Thus, we have N equations for N unknowns: $I_1, I_2, \ldots, I_N$. For given initial conditions, this system has, generally speaking, a single-valued solution.

Case of Two Conductors. As an example, we shall consider the system of equations for two conductors ($N = 2$)

$$I_1 R_1 = \mathcal{E}_1^{\text{app}} - \left(L_{11} \frac{dI_1}{dt} + L_{12} \frac{dI_2}{dt}\right) \tag{24.11}$$

$$I_2 R_2 = \mathcal{E}_2^{\text{app}} - \left(L_{21} \frac{dI_1}{dt} + L_{22} \frac{dI_2}{dt}\right) \tag{24.12}$$

The term $-L_{11}\, dI_1/dt$ on the right-hand side of (24.11) takes into account the self-induction emf in the first conductor, while the term $-L_{12}\, dI_2/dt$ takes into account the induced emf which arises in the first conductor due to a change in the magnetic flux set up by the second conductor. The terms on the right-hand side of the second equation have analogous meanings. Let the first conductor be the primary coil of a transformer, to which a voltage

$$\mathcal{E}_1^{\text{app}} = \mathcal{E}_{10}^{\text{app}} e^{i\omega t} \tag{24.13}$$

is applied, and let the second conductor be the secondary coil, in which there is no applied emf

$$\mathcal{E}_2^{\text{app}} = 0$$

We seek the solution of the system of equations in the form

$$I_i = I_{i0} e^{i\omega t} \tag{24.14}$$

After substituting (24.14) in (24.11) and (24.12) and cancelling the common factor $e^{i\omega t}$, we obtain

$$I_{10}R_1 = \mathcal{E}_{10}^{\text{app}} - i\omega(L_{11}I_{10} + L_{12}I_{20}) \tag{24.15}$$

$$I_{20}R_2 = -i\omega(L_{21}I_{10} + L_{22}I_{20}) \tag{24.16}$$

From equation (24.16), it follows that

$$\frac{I_{20}}{I_{10}} = \frac{-i\omega L_{21}}{R_2 + i\omega L_{22}} \tag{24.17}$$

Taking moduli, we obtain

$$\frac{|I_{20}|}{|I_{10}|} = \frac{\omega L_{21}}{\sqrt{R_2{}^2 + \omega^2 L_{22}{}^2}} \tag{24.18}$$

Generally, the ohmic resistance R_2 of the secondary coil is much less than its reactance

$$R \ll \omega L_{22} \tag{24.19}$$

and equation (24.18) may be rewritten

$$\frac{|I_{20}|}{|I_{10}|} \approx \frac{L_{21}}{L_{22}} \tag{24.20}$$

It is clear that if the coefficient of self-induction of the secondary coil is less than the coefficient of mutual induction of the primary and secondary coils, then the current in the secondary coil is greater than that in the primary coil, i.e., the system of conductors acts as a transformer.

§25. Electric Circuit with Capacitance and Inductance

A simplified electrical circuit with capacitance and inductance is given in Fig. 34.

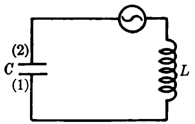


Fig. 34

Equation for Dealing with an Electric Circuit. We multiply both sides of the equation

$$\frac{\mathbf{j}}{\lambda} = \mathbf{E} + \mathbf{E}^{\text{app}} \tag{25.1}$$

by an element of length of the conductor $d\mathbf{l}$ and integrate along the

conductor from one plate of the capacitor to the other, i.e., from point 1 to point 2 (Fig. 34)

$$\int_{(1)}^{(2)} \frac{\mathbf{j}\cdot d\mathbf{l}}{\lambda} = \int_{(1)}^{(2)} \mathbf{E}\cdot d\mathbf{l} + \int_{(1)}^{(2)} \mathbf{E}^{\text{app}}\cdot d\mathbf{l} \tag{25.2}$$

The left-hand integrand is transformed similarly to (24.3), and the right-hand integrand similarly to (24.5) and (24.6). We thus obtain

$$\int_{(1)}^{(2)} I\, dR = \mathcal{E}^{\text{app}} - \int_{(1)}^{(2)} \text{grad}\, \varphi\cdot d\mathbf{l} - \frac{d}{dt}\int_{(1)}^{(2)} \mathbf{A}\cdot d\mathbf{l} \tag{25.3}$$

The integral

$$\int_{(1)}^{(2)} \text{grad}\, \varphi\cdot d\mathbf{l} = \int_{(1)}^{(2)} d\varphi = \varphi_2 - \varphi_1 \tag{25.4}$$

gives the potential difference between the plates of the capacitor. The integral

$$\int_{(1)}^{(2)} \mathbf{A}\cdot d\mathbf{l} \tag{25.4'}$$

is evaluated approximately, as follows. Since $\mathbf{A}$ is a continuous function, and the distance between the plates of the condenser is much less than the length of the conductor along which (25.4′) is evaluated, we may conclude that

$$\int_{(1)}^{(2)} \mathbf{A}\cdot d\mathbf{l} \approx \oint_{L_1} \mathbf{A}\cdot d\mathbf{l} = \Phi$$

where L_1 is a closed contour consisting of the conductor and the short path between the plates of the capacitor, and Φ is the magnetic induction flux through the surface subtended by L_1.

The left-hand side of (25.3) is evaluated in a similar manner to (24.4), and (25.3) may be written in the form

$$IR = \mathcal{E}^{\text{app}} - (\varphi_2 - \varphi_1) - \frac{d\Phi}{dt} \tag{25.5}$$

We shall use

$$\Phi = LI \tag{25.6}$$

where L is the coefficient of self-induction of the contour under consideration. We then allow for the fact that the potential difference $\varphi_2 - \varphi_1$ between the capacitor plates is related to the charge q on the plates by the equation

$$\varphi_2 - \varphi_1 = \frac{q}{C} \tag{25.7}$$

where C is the capacitance of the capacitor. Substituting (25.6) and (25.7) in (25.5), we obtain

$$L\frac{dI}{dt} + RI + \frac{1}{C}q = \mathcal{E}^{\text{app}} \tag{25.8}$$

Differentiating both sides of (25.8) with respect to time, and remembering that

$$\frac{dq}{dt} = I \tag{25.9}$$

we obtain the equation for an electric circuit with capacitance C and inductance L in the form

$$L\frac{d^2I}{dt^2} + R\frac{dI}{dt} + \frac{1}{C}I = \frac{d}{dt}\mathcal{E}^{\text{app}} \tag{25.10}$$

Thus, the solution of problems connected with electric circuits of this form reduces to the solution of an ordinary differential equation (25.10) with constant coefficients. This may generally be solved without difficulty. Two examples are considered below.

Switching On and Off a Constant EMF in a Circuit. At the instant of time $t = 0$, let a constant emf $\mathcal{E}_0^{\text{app}}$ be switched into a circuit. We shall suppose that there is no capacitor in the circuit. We wish to find the law of current rise in the circuit, i.e., the function $I = I(t)$.

To solve the problem, it is convenient to use equation (25.8), which, for $t > 0$, has the form

$$L\frac{dI}{dt} + RI = \mathcal{E}_0^{\text{app}} \tag{25.11}$$

The initial condition may be written

$$I(0) = 0 \tag{25.12}$$

The general solution of equation (25.11) has the form

$$I = \frac{\mathcal{E}_0^{\text{app}}}{R} + ae^{-\frac{R}{L}t} \tag{25.13}$$

where a is an arbitrary constant, found from the initial condition (25.12). This initial condition is written, on the basis of (25.13), in the form

$$\frac{\mathcal{E}_0^{\text{app}}}{R} + a = 0 \qquad a = -\frac{\mathcal{E}_0^{\text{app}}}{R} \tag{25.14}$$

Thus, the solution of the problem is given by

$$I(t) = \frac{\mathcal{E}_0^{\text{app}}}{R}(1 - e^{-\frac{R}{L}t}) \qquad t \geqslant 0 \tag{25.15}$$

The problem of the switching off of a constant emf is solved in a completely analogous manner. The law of current decay is given by

$$I(t) = \frac{\mathcal{E}_0^{\text{app}}}{R} e^{-\frac{R}{L}t} \qquad t \geqslant 0 \tag{25.16}$$

Thus, owing to self-induction, when a constant emf is switched off, the current does not die away to zero instantaneously, but over an interval of time. The interval during which the current decreases e times is called the *relaxation time*. From (25.16) it is clear that the relaxation time is

$$t_r = \frac{L}{R} \tag{25.17}$$

Oscillating Circuit. We shall assume that the ohmic resistance of a circuit may be ignored ($R = 0$), and there is no applied emf ($\mathcal{E}^{\text{app}} = 0$). Then, equation (25.10) becomes

$$\frac{d^2I}{dt^2} + \frac{1}{LC} I = 0 \tag{25.18}$$

The general solution of this equation is

$$I = A \sin \omega t + B \cos \omega t \tag{25.19}$$

where the circular frequency ω is

$$\omega = \frac{2\pi}{T} = \frac{1}{\sqrt{LC}} \tag{25.20}$$

The period of oscillations T in such a circuit is

$$T = 2\pi\sqrt{LC} \tag{25.21}$$

These oscillations may be excited by the electromagnetic induction. The presence of an ohmic resistance leads to a gradual damping of the oscillations.

General Case. We shall now consider equation (25.10), assuming the presence of capacitance, resistance and inductance. If the applied emf is a periodic function of frequency ω

$$\mathcal{E}^{\text{app}} = \mathcal{E}_0^{\text{app}} e^{-i\omega t} \tag{25.22}$$

then the solution of equation (25.10) must be sought in the form

$$I = I_0 e^{-i\omega t} \tag{25.23}$$

Substituting (25.23) in (25.1) and differentiating, we find the following relationship

$$\left(-L\omega^2 - Ri\omega + \frac{1}{C}\right) I = -i\omega\mathcal{E}^{\text{app}} \tag{25.24}$$

This relationship may be written in the form of Ohm's law

$$ZI = \mathcal{E}^{\text{app}} \tag{25.25}$$

where

$$Z = R - i\left(\omega L - \frac{1}{\omega C}\right) \tag{25.26}$$

is called the *impedance*. Writing (25.25) in the form

$$I = \frac{\mathcal{E}^{\text{app}}}{Z} \tag{25.27}$$

and separating the real part, we obtain the following expression for the current

$$I(t) = \frac{\mathcal{E}_0^{\text{app}} \cos(\omega t - \alpha)}{\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}}$$

$$\tan \alpha = \left(\omega L - \frac{1}{\omega C}\right)\frac{1}{R} \tag{25.28}$$

This current is produced by the externally applied voltage

$$\mathcal{E}^{\text{app}} = \mathcal{E}_0^{\text{app}} \cos \omega t \tag{25.29}$$

Thus, the current suffers a phase shift with respect to the applied voltage, and an inductive resistance appears in addition to the ohmic resistance.

If the applied emf is switched off, then the current in the circuit oscillates with a complex frequency defined, on the basis of (25.5), by the condition $Z = 0$. Consequently

$$\omega = -i\frac{R}{2L} \pm \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2} \tag{25.30}$$

If the discriminant of this equation is negative, $1/LC < (R/2L)^2$, then the complex frequency is purely imaginary, and, consequently, the factor $e^{-i\omega t}$ is not a periodic function. In this case, an aperiodic damped discharge takes place.

If the discriminant is positive, $1/LC > (R/2L)^2$, then damped oscillations are set up with the frequency

$$\omega_{\text{osc}} = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2} \tag{25.31}$$

and the rate of decrease of the amplitude is given by

$$e^{-\frac{R}{2L}t} \tag{25.32}$$

The quantity $R/2L$ is called the *damping decrement* of these oscillations. If we ignore the small ohmic resistance in (25.31), putting $R = 0$, this

formula becomes Thompson's formula (25.20) for the frequency of free oscillations.

§26. Induction of Currents in Moving Conductors

In considering the phenomenon of Faraday's electromagnetic induction in §6, we have assumed that the circuit L, in which the induced emf was calculated, was stationary, and that the change in the magnetic induction flux, Φ, was caused by the variation of the magnetic field with time.

We shall now consider a closed linear conductor L, which moves in an arbitrary manner in an external magnetic field $\mathbf{B}$. Deformation of the shape of the conductor may also occur.

If an element of the conductor $d\mathbf{l}$ moves with a velocity $\mathbf{v}$ in the magnetic field $\mathbf{B}$, then the Lorentz force acts on every free electron in the element

$$\mathbf{F} = e\mathbf{v} \times \mathbf{B} \tag{26.1}$$

and this force sets up an ordered motion of electrons, i.e., an electric current. From this point of view, we may say that some effective electric field

$$\mathbf{E}_{\text{eff}} = \mathbf{v} \times \mathbf{B} \tag{26.2}$$

is set up in the conductor, and this produces an induced emf. The magnitude of this emf is

$$\mathcal{E}^{\text{ind}} = \oint_L \mathbf{E}_{\text{eff}} \cdot d\mathbf{l} = \oint_L (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} \tag{26.3}$$

Let $\mathbf{r}$ be the radius vector of $d\mathbf{l}$ (Fig. 35). We shall consider two positions of L: initial, and after an interval of time δt. During δt, $d\mathbf{l}$ moves a distance $\delta\mathbf{r}$, and

$$\mathbf{v} = \lim_{\delta t \to 0} \frac{\delta \mathbf{r}}{\delta t} \tag{26.4}$$

Hence, equation (26.3) may be put in the form

$$\mathcal{E}^{\text{ind}} = \lim_{\delta t \to 0} \frac{1}{\delta t} \oint_L (\delta\mathbf{r} \times \mathbf{B}) \cdot d\mathbf{l} \tag{26.5}$$

Using the cyclic rule for a scalar triple product, we have

$$(\delta\mathbf{r} \times \mathbf{B}) \cdot d\mathbf{l} = (d\mathbf{l} \times \delta\mathbf{r}) \cdot \mathbf{B} \tag{26.6}$$

Also

$$d\mathbf{l} \times \delta\mathbf{r} = -d\mathbf{S}_{\text{lat}} \tag{26.7}$$

where $d\mathbf{S}_{\text{lat}}$ is the area swept by the element $d\mathbf{l}$ when it is displaced by $\delta\mathbf{r}$. The vector $d\mathbf{S}_{\text{lat}}$ lies along the outward normal to the volume enclosed

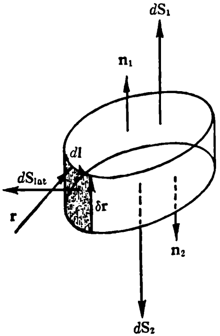


Fig. 35

by S_1, S_2, and S_{lat} in Fig. 35. The flux of **B** through any closed surface equals zero

$$\oint_S \mathbf{B}\cdot d\mathbf{S} = \int_{S_1} \mathbf{B}\cdot d\mathbf{S} + \int_{S_2} \mathbf{B}\cdot d\mathbf{S} + \int_{S_{lat}} \mathbf{B}\cdot d\mathbf{S} = 0 \qquad (26.8)$$

Consequently

$$-\int_{S_{lat}} \mathbf{B}\cdot d\mathbf{S} = \int_{S_{lat}} \mathbf{B}\cdot(d\mathbf{l}\times\delta\mathbf{r}) = \int_{S_1} \mathbf{B}\cdot d\mathbf{S} + \int_{S_2} \mathbf{B}\cdot d\mathbf{S} \qquad (26.9)$$

The positive normal to the surface subtended by the moving contour must be chosen so as to make a right-hand screw system with the direction of describing the contour. Taking the latter direction as shown in Fig. 35, $\mathbf{n}_1$ is the required positive normal. Therefore, the integrals on the right-hand side of (26.9) may be written

$$\int_{S_1} \mathbf{B}\cdot d\mathbf{S} = \Phi(t) \qquad \int_{S_2} \mathbf{B}\cdot d\mathbf{S} = -\Phi(t+\delta t) \qquad (26.10)$$

The minus sign in the second equation is due to the outward normal $\mathbf{n}_2$ being opposite in direction to the positive direction. Thus, (26.9) may be written

$$\int_{S_{lat}} (\mathbf{B}\cdot d\mathbf{l}\times\delta\mathbf{r}) = -[\Phi(t+\delta t) - \Phi(t)] = -\delta\Phi \qquad (26.11)$$

where $\delta\Phi$ is the change in the magnetic induction flux through the surface subtended by L, due to the motion and deformation of the contour.

Substituting (26.11) in (26.5), we find

$$\mathcal{E}^{\text{ind}} = -\lim_{\delta t \to 0} \frac{\delta \Phi}{\delta t} = -\frac{d\Phi}{dt} \tag{26.12}$$

This expression is outwardly identical with equation (6.1), but its meaning is essentially different: equation (6.1) gives the change in the magnetic induction flux of a stationary conductor, produced by a change in the magnetic field, while (26.12) gives the change in the magnetic induction flux produced by the motion and deformation of the conductor in a magnetic field that is constant with respect to time. Consequently, the expression

$$\mathcal{E}^{\text{ind}} = -\frac{d\Phi}{dt} \tag{26.13}$$

applies, whatever the reasons for the change of Φ.

§27. Skin Effect

Nature of the Phenomenon. A steady current is distributed uniformly over the cross section of a conductor. In the case of alternating currents the picture is different; the current density is higher near the surface of the conductor and lower near its center. This phenomenon is called the *skin effect*, and is caused by the electromagnetic interaction between current elements.

Consider an infinite right circular cylindrical conductor, along which flows an alternating current (Fig. 36). Let us assume that the current is flowing, at some instant of time, in the direction indicated by the arrow in Fig. 36, and let the strength of the current be increasing. The ines of force of the magnetic field of the current are concentric circles with center on the axis of the conductor. As the current increases, the shape of these lines of force remains the same, but the intensity of the field increases at every point. Consequently, as the current increases, the intensity of the magnetic field at every point still keeps the same direction, but increases in magnitude. Hence, the derivative $\partial \mathbf{B}/\partial t$ lies along the tangent to the corresponding line of force of the magnetic field. By the law of magnetic induction, the changing magnetic field produces an induced electric field, and the lines of force of this field lie in a plane perpendicular to $\partial \mathbf{B}/\partial t$. The positive direction of describing the lines of force of the electric field and the direction of $\partial \mathbf{B}/\partial t$ form a left-hand screw system. As is immediately evident from Fig. 36, the direction of the induced field $\mathbf{E}$ is such that it makes the current stronger near the surface and weaker inside the con-

ductor. We have considered the case of increasing current. It is not difficult to verify that in the case of decreasing current, the direction of the induced field is also such that the current density increases outwards from the center of the conductor to the surface.

Elementary Theory of the Skin Effect. To simplify the calculation, we shall consider the case of an infinite homogeneous conductor occupying the half-space $y \geqslant 0$. A current flows in the direction of the x axis, parallel

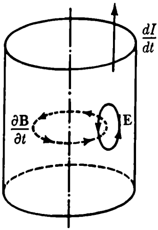


Fig. 36

to the surface of the conductor (the x, z plane). The initial equations have the form

$$\text{curl}\, \mathbf{H} = \mathbf{j} = \lambda \mathbf{E} \tag{27.1}$$

$$\text{curl}\, \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{27.2}$$

where we have allowed for

$$\mathbf{j} = \lambda \mathbf{E} \tag{27.3}$$

Differentiating both sides of (27.1) with respect to time, and eliminating $\partial \mathbf{H}/\partial t$ from the left-hand side by means of (27.2), we find

$$-\frac{1}{\mu}\, \text{curl curl}\, \mathbf{E} = \lambda \frac{\partial \mathbf{E}}{\partial t} \tag{27.4}$$

Since

$$\text{curl curl}\, \mathbf{E} = \text{grad div}\, \mathbf{E} - \Delta \mathbf{E}$$

and div $\mathbf{E} = 0$ because there are no free charges in a homogeneous conductor, we obtain from (27.4) the following equation for $\mathbf{E}$

$$\Delta \mathbf{E} = \lambda\mu \frac{\partial \mathbf{E}}{\partial t} \tag{27.5}$$

The equation for **H** is deduced in a completely analogous manner, giving

$$\Delta \mathbf{H} = \lambda\mu \frac{\partial \mathbf{H}}{\partial t} \tag{27.6}$$

We shall consider the case when the current flows along the x axis and

$$j_x = j_x(y, t) \qquad j_y = j_z = 0 \tag{27.7}$$

Then, on the basis of (27.3), we have

$$E_x = E_x(y, t) \qquad E_y = E_z = 0 \tag{27.8}$$

and, therefore, (27.5) takes the form

$$\frac{\partial^2 E_x}{\partial y^2} = \lambda\mu \frac{\partial E_x}{\partial t} \tag{27.9}$$

If ω is the frequency of the alternating current, then the solution of (27.9) must be sought in the form

$$E_x(y, t) = E_{x0}(y)e^{i\omega t} \tag{27.10}$$

Substituting this expression for $E_x(y, t)$ in (27.9), and simplifying after differentiation with respect to time, by cancelling the factor $e^{i\omega t}$, we obtain

$$\frac{d^2 E_{x0}}{dy^2} = 2ip^2 E_{x0} \qquad p^2 = \frac{1}{2}\mu\lambda\omega \tag{27.11}$$

The general solution of this equation has the form

$$E_{x0}(y) = A_0 e^{ky} + B_0 e^{-ky} \tag{27.12}$$

where

$$k^2 = 2ip^2 \qquad k = p\sqrt{2i} = p(1 + i) \tag{27.13}$$

Thus, we obtain

$$E_{x0} = A_0 e^{py} e^{ipy} + B_0 e^{-py} e^{-ipy} \tag{27.14}$$

The first term of (27.14) increases without limit as $y \to \infty$, which is physically meaningless. Hence, we must put $A_0 = 0$. Taking (27.10) into account, we finally obtain

$$E_x(y, t) = e^{-py} B_0 e^{i(\omega t - py)} \tag{27.15}$$

or, taking, for example, the real part, we have

$$E_x = e^{-py} B_0 \cos(\omega t - py) \tag{27.16}$$

Using (27.3) and (27.16), we can rewrite the expression for the current density in the following form

$$j_x = e^{-py} j_0 \cos(\omega t - py) \tag{27.17}$$

where j_0 is the amplitude of the current density on the surface of the conductor. Thus, the current density decreases with distance from the surface of the conductor, and the rate of decrease is described by the exponential factor e^{-py}. At a distance

$$\Delta = \frac{1}{p} \tag{27.18}$$

from the surface of the conductor, the current density decreases by a factor e. Thus, for practical purposes, we may assume that all the current is concentrated in a surface layer of thickness Δ. Using equation (27.11), equation (27.18) may be put in the form

$$\Delta = \sqrt{\frac{T}{\pi\mu\lambda}} \tag{27.19}$$

where T is the period of the oscillations. Thus, as the frequency of the alternating current increases, the skin effect also increases and the current is concentrated into a still thinner layer near the surface of the conductor. The skin effect also increases on increase of the conductivity of the conductor, i.e., on decrease of the resistivity. Let us consider the order of Δ. For an order of magnitude estimate for metals, we may take $\mu = \mu_0$, $\lambda \approx 10^7$ ohm^{-1} cm^{-1}. Then for $T \approx 10^{-3}$ sec, we obtain

$$\Delta = \sqrt{\frac{10^{-3}}{\pi\, 4\pi\, 10^{-7}\, 10^7}} \approx 0.5 \times 10^{-2}\ \mathrm{m} = 5\ \mathrm{mm}$$

If the frequency is increased by a factor of 100, then

$$\Delta = 0.5\ \mathrm{mm}$$

Thus, when the period $T = 10^{-5}$ sec, which corresponds to a wavelength $l = cT = 3$ km, for all practical purposes the entire current flows in a layer half a millimeter thick. These estimates show that at sufficiently high frequencies, the skin effect leads to a very considerable redistribution of the current over the cross section of the conductor.

The calculation of the skin effect in a cylindrical conductor may be carried out by a similar method. No new ideas are involved, but the argument is more cumbersome, and, hence, it is not given here.

Dependence of the Ohmic Resistance on the Frequency. As is well known, the ohmic resistance of a conductor is inversely proportional to the area of the cross section. Owing to the skin effect, the entire current becomes concentrated close to the surface of a cylindrical conductor. This is equivalent to the entire current flowing in a hollow cylindrical conductor, and the resistance of this conductor becomes equal to the resistance of a hollow cylinder of corresponding thickness, i.e., it becomes greater than

the resistance of the solid conductor. As the frequency increases, the thickness of the cylindrical layer carrying the current decreases and, hence, the resistance increases. Thus, except at low frequencies, the principal carrier of the current is the surface layer of the conductor. Hence, we can make the inner part of the cylindrical conductor of some poorly conducting material with an outer coating of metal of high conductivity, which is usually more expensive. The principal purpose of the inner core is then to impart strength to the conductor. In general, the use of such conductors proves, in many cases, to be economically and technically advantageous.

Dependence of the Coefficient of Self-Induction on the Frequency. As shown in §21, the energy of a magnetic field is related to the coefficient of self-induction L and the current I by the relationship

$$W = \frac{1}{2} LI^2 \tag{27.20}$$

If the current flows in a hollow cylinder, then the magnetic field inside the cavity is equal to zero, and the field outside the cylinder is equal to that obtained for the same current flowing in a solid cylindrical conductor. Hence, the energy of the magnetic field set up by a current flowing in a hollow cylindrical conductor is less than the energy of the magnetic field set up by a current of equal strength flowing in a solid cylindrical conductor. In the case of the skin effect, the current is concentrated close to the surface of the conductor. Consequently, the field outside the conductor is the same as in the absence of the skin effect, but the field inside the conductor is less. Therefore, due to the skin effect, the energy of the magnetic field decreases. Since the total current flowing in the conductor remains the same, then, by equation (27.20), this drop in the energy can be due only to a decrease in the coefficient of self-induction L. Thus, the coefficient of self-induction depends on the frequency, and decreases as the frequency of the alternating current increases.

PROBLEMS

1 A horizontal metal rod rotates about a vertical axis located at a distance $1/k$ of its length from one of its ends, making N revolutions per second. The length of the rod is l. Determine the potential difference between the ends of the rod if it is revolving in a vertical homogeneous magnetic field of induction B. *Example:* $k = 3$, $l = 120$ cm, $N = 6\ \text{sec}^{-1}$, $B = 10^{-2}$ tesla.
Hint: Use the method of §26.

Answer: $V = \pi N l^2 \dfrac{k-2}{k}$

2 A conducting frame of resistance R is rotating in a constant magnetic field of induction B about an axis perpendicular to the field. Determine what quantity of electricity passes through a galvanometer connected in series with the frame, if initially the plane of the frame was perpendicular to the field and the plane has turned through 90°. The area of the frame is S.

Solution:

$$\mathcal{E}^{\text{ind}} = -\frac{d\Phi}{dt} \qquad I = \frac{\mathcal{E}^{\text{ind}}}{R}$$

$$q = \int I\,dt = -\frac{1}{R}\int \frac{d\Phi}{dt}\,dt = \frac{\Delta\Phi}{R}$$

3 A closed circuit in the form of a frame of area S rotates uniformly in a homogeneous magnetic field B, making N revolutions per second. Determine the maximum emf in the circuit.

Consider the numerical example: $B = 1$ tesla, $S = 1000\ \text{cm}^2$, $N = 10\ \text{sec}^{-1}$.

Answer: $\mathcal{E}^{\text{ind}}_{\text{max}} = 2\pi BSN = 6.28$ V

4 An oscillating circuit consists of a coil of inductance L, and a plane capacitor with plates of area S separated by a dielectric of thickness d and permittivity ϵ. Determine the period of the oscillations of this circuit. The resistance is negligibly small.

Consider the numerical example: $L = 0.1$ henry, $S = 500\ \text{cm}^2$, $d = 1$ mm, $\epsilon = 2\epsilon_0$.

Answer: $T = 2\pi\sqrt{L\dfrac{S\epsilon}{d}} \approx 2.2 \times 10^{-2}$ sec

5 A voltage is applied to a coil of resistance $R = 5$ ohms and inductance $L = 100$ millihenries (mh). How long will it take for the current in the coil to reach half its steady value?

Answer: $t = \dfrac{L}{R}\ln 2 \approx 0.014$ sec

6 A magnetic field established between the circular poles of an electromagnet connected to a 1000 cycles/sec (cps) a.c. supply varies sinusoidally with time. The amplitude of the induction is $B_0 = 5 \times 10^3$ gauss. Assuming the magnetic field to be homogeneous, determine the maximum value of the field intensity in the gap between the poles at a distance $r = 0.5$ m from the center.
Hint: Use the law of electromagnetic induction.

Answer: $E = \dfrac{1}{2}B_0\omega r = 7.8 \times 10^2\ \text{V/m} = 7.8\ \text{V/cm}$

7 A solenoid, closed on itself, of radius b and n turns, is rotating with angular velocity ω about the diameter of one of its turns in a homogeneous magnetic field H. The axis of rotation is perpendicular to the magnetic field. The resistance

and coefficient of self-induction of the solenoid are R and L, respectively. Determine the current flowing in the solenoid as a function of time.

$$\textit{Answer:}\quad I = \mu_0\pi b^2 nH\omega(R^2 + \omega^2L^2)^{-1/2}\sin(\omega t + \varphi_0)$$

where φ_0 is the phase at $t = 0$.

8 A superconducting ring, which can move only in the vertical direction, is lying on a table above a loop of wire. A current I begins to flow through the loop. As a result, the superconducting ring rises. The coefficient of mutual induction of the loop and the ring when it has risen a height x is $L_{12}(x)$. The coefficient of self-induction of the ring is L_{11}, the mass of the ring is m, and the acceleration due to gravity is g. Determine the height to which the ring will rise.

$$\textit{Answer:}\quad h = \frac{1}{2}\frac{I^2}{mg}\frac{1}{L_{11}}\{[L_{12}(0)]^2 - [L_{12}(h)]^2\}$$

9 A current $I_0 \sin\omega t$ flows in a coil A_1. A corresponding current is induced in a coil A_2. The coefficients of self-induction and mutual induction of the coils are L_1, L_2, and L_{12}, and the resistance of A_2 is R_2. Let q_i be some coordinate describing the position of A_2. Find the generalized mean force F_i associated with changes in the generalized coordinate q_i.

$$\textit{Answer:}\quad F_i = -\frac{1}{2}\frac{I_0^2\omega^2L_2L_{12}}{R^2 + \omega^2L_2^2}\frac{\partial L_{12}}{\partial q_i}$$

CHAPTER 5

Generation of Electromagnetic Waves

§28. General Equations

In our discussion of quasi-static electromagnetic fields, the interrelationship between electric and magnetic fields has not been allowed for in full: the excitation of an electric field by a changing magnetic field (electromagnetic induction) is taken into account, but not the excitation of a magnetic field by a changing electric field (displacement currents). A full discussion of the interdependence of electric and magnetic fields reveals the physical existence of electromagnetic waves.

The electric field in electrostatics and the magnetic field in magnetostatics always exist in association with electric charges and electric currents, respectively; they do not exist apart from the charges and currents, they do not "break away" from their sources. Only in the case of electromagnetic waves does the electromagnetic field attain complete independence, it "breaks away" from the currents and charges that generate it, and exists independent of what later happens to these currents and charges.

Vector and Scalar Potentials. In the following discussion, no restrictions are placed on the rate of change of the fields. We start from the full system of Maxwell's equations (8.1). In contrast to the case of quasi-static fields, the displacement current is now included. However, the inclusion of the displacement current does not involve any change in the relationships connecting the intensities of the fields with the potentials. This is immediately evident from the deduction of equations (23.12) and (23.18), and, therefore, as in the quasi-static case, we may write

$$\mathbf{B} = \operatorname{curl} \mathbf{A} \tag{28.1}$$

$$\mathbf{E} = -\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} \tag{28.2}$$

Nonuniqueness of the Potentials. Gauge Transformation. It is not possible, given E and B, to determine the potentials uniquely from equations (28.1) and (28.2). Let $\chi(\mathbf{r}, t)$ be some arbitrary continuous function of the coordinates and time. Let φ and A be the potentials which, by (28.1) and (28.2), describe some electromagnetic field **E** and **B**. We assert that the potentials

$$\mathbf{A}' = \mathbf{A} + \operatorname{grad} \chi \qquad \varphi' = \varphi - \frac{\partial \chi}{\partial t} \tag{28.3}$$

describe the same field **E** and **B**. To prove this, we find the electromagnetic field $\mathbf{E}'$ and $\mathbf{B}'$ described by the potentials $\mathbf{A}'$ and φ' from (28.3)

$$\mathbf{B}' = \operatorname{curl} \mathbf{A}' = \operatorname{curl} (\mathbf{A} + \operatorname{grad} \chi) = \operatorname{curl} \mathbf{A} + \operatorname{curl} \operatorname{grad} \chi = \operatorname{curl} \mathbf{A} = \mathbf{B}$$

Similarly

$$\mathbf{E}' = -\operatorname{grad} \varphi' - \frac{\partial \mathbf{A}'}{\partial t} = -\operatorname{grad} \left(\varphi - \frac{\partial \chi}{\partial t}\right) - \frac{\partial}{\partial t} (\mathbf{A} + \operatorname{grad} \chi)$$

$$= -\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} = \mathbf{E}$$

since the order of differentiation with respect to independent variables may be reversed.

Thus, the potentials A and φ describe the same electromagnetic field as do φ' and $\mathbf{A}'$, which are related to φ and A by equation (28.3), where χ is an arbitrary function. The transformation (28.3) is called the *gauge* or *gradient transformation*. Using this arbitrariness in the choice of the potential, we may choose the potentials so that they observe some given additional conditions. As an example of such an additional condition in electrodynamics, we may use Lorentz's condition

$$\operatorname{div} \mathbf{A} + \epsilon\mu \frac{\partial \varphi}{\partial t} = 0 \tag{28.4}$$

This additional condition is chosen in such a form as to make the equations for the potentials as simple as possible, as will be seen from the following. In the general case, Lorentz's condition is not invariant with respect to the gauge transformation (28.3). In fact, if the potentials φ and **A** satisfy the condition (28.4), then the potentials φ' and $\mathbf{A}'$, defined by (28.3), satisfy the condition

$$\operatorname{div} \mathbf{A}' + \epsilon\mu \frac{\partial \varphi'}{\partial t} = \Delta\chi - \epsilon\mu \frac{\partial^2 \chi}{\partial t^2} \tag{28.5}$$

Since the function χ is arbitrary, the right-hand side of (28.5) is not, in general, equal to zero. But, if χ is chosen to be a function satisfying the wave equation

$$\Delta\chi - \epsilon\mu \frac{\partial^2 \chi}{\partial t^2} = 0 \tag{28.6}$$

then equation (28.5) is transformed into Lorentz's condition (28.4). Hence, if it is necessary to use the gauge transformation (28.3) to pass from some potentials φ, $\mathbf{A}$, which satisfy Lorentz's condition, to other potentials φ' and $\mathbf{A}'$, which also satisfy Lorentz's condition, then χ must be chosen in such a way that it satisfies the wave equation.

Equation for the Vector Potential. We shall consider a homogeneous medium: $\mu = \text{const}$, $\epsilon = \text{const}$. Substituting equations (28.1) and (28.2) in Maxwell's equation

$$\operatorname{curl} \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}$$

we obtain

$$\operatorname{curl} \operatorname{curl} \mathbf{A} = \mu \mathbf{j} + \epsilon\mu \frac{\partial}{\partial t}\left(-\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t}\right)$$

The left-hand side of this equation may be transformed by using the formula (A.8) in Appendix 1. Thus, we obtain

$$\Delta \mathbf{A} - \epsilon\mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{j} + \operatorname{grad}\left(\operatorname{div} \mathbf{A} + \epsilon\mu \frac{\partial \varphi}{\partial t}\right)$$

To simplify this equation as far as possible, we use the fact that the choice of the potentials is arbitrary, and equate the final term on the right-hand side to zero. This also gives the Lorentz condition (28.4). Thus, finally, we obtain the equation for the vector potential in the form

$$\Delta \mathbf{A} - \epsilon\mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{j} \tag{28.7}$$

where $\mathbf{A}$ and φ satisfy the Lorentz condition (28.4).

Equation for the Scalar Potential. We shall substitute the expression for $\mathbf{E}$ from (28.2) in Maxwell's equation

$$\operatorname{div} \mathbf{D} = \rho$$

We thus obtain, for a homogeneous medium ($\mu = \text{const}$ and $\epsilon = \text{const}$)

$$\operatorname{div}\left(-\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t}\right) = \frac{\rho}{\epsilon}$$

or

$$\Delta\varphi + \frac{\partial}{\partial t} \operatorname{div} \mathbf{A} = -\frac{\rho}{\epsilon}$$

Since φ and A satisfy Lorentz's condition (28.4), we finally obtain the expression for the scalar potential in the form

$$\Delta\varphi - \epsilon\mu \frac{\partial^2 \varphi}{\partial t^2} = -\frac{\rho}{\epsilon} \tag{28.8}$$

Fundamental Data on the Solution of D'Alembert's Equation. The equations for the scalar and vector potentials have the following form

$$\Delta\Phi - \frac{1}{v^2}\frac{\partial^2 \Phi}{\partial t^2} = -f(x, y, z, t) \tag{28.9}$$

This is *D'Alembert's equation.*

If $f = 0$, then we obtain D'Alembert's homogeneous equation

$$\Delta\Phi - \frac{1}{v^2}\frac{\partial^2 \Phi}{\partial t^2} = 0 \tag{28.10}$$

In the theory of partial differential equations, it is proved that equation (28.10) describes a wave propagated with a velocity v. We shall consider the one-dimensional case as an example. The equation in this case takes the form

$$\frac{\partial^2 \Phi}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 \Phi}{\partial t^2} = 0 \tag{28.11}$$

It is easy to see that this equation is satisfied by an arbitrary function ψ with the argument $t - x/v$ or $t + x/v$. For example

$$\Phi = \psi\left(t - \frac{x}{v}\right) \tag{28.12}$$

We have

$$\begin{aligned} &\frac{\partial \psi}{\partial x} = \psi'\left(-\frac{1}{v}\right) \qquad \frac{\partial^2 \psi}{\partial x^2} = \psi''\frac{1}{v^2} \\ &\frac{\partial \psi}{\partial t} = \psi' \qquad \frac{\partial^2 \psi}{\partial t^2} = \psi'' \end{aligned} \tag{28.13}$$

where the prime denotes differentiation of the function with respect to its argument $t - x/v$. Direct substitution of the function (28.12) in (28.11), using (28.13), shows that this function does, in fact, satisfy the equation. Transforming the argument $t - x/v$

$$t - \frac{x}{v} = t + \Delta t - \frac{x + \Delta x}{v} \qquad \Delta t = \frac{\Delta x}{v}$$

we see that the value of the function ψ at time $t + \Delta t$ at the point $x + \Delta x$ is the same as at time t and point x. This means that we have a wave traveling with velocity v in the positive direction along the x axis. Similarly, it may be shown that the arbitrary function

$$\Phi = \psi\left(t + \frac{x}{v}\right) \tag{28.14}$$

is also a solution of (28.11) and describes a wave traveling with velocity v along the x axis in the negative direction. Thus, equation (28.10) describes waves traveling in three-dimensional space with velocity v. We have verified this assertion for a special case, but in the theory of differential equations it is proved in the general case.

From a comparison of (28.7) and (28.8) with (28.9) and (28.10) it is clear that the velocity of propagation of electromagnetic waves is

$$v = \frac{1}{\sqrt{\epsilon\mu}} = \frac{c}{\sqrt{\epsilon'\mu'}} \tag{28.15}$$

where $c = 1/\sqrt{\epsilon_0\mu_0}$ is the velocity of light and $\epsilon' = \epsilon/\epsilon_0$, $\mu' = \mu/\mu_0$ are the relative permittivity and permeability of the medium.

In the theory of differential equations, it is proved that if the function f on the right-hand side of equation (28.9) is different from zero in a finite region of space, then the solution of the equation for all space will have the form

$$\Phi(x, y, z, t) = \frac{1}{4\pi}\int_V \frac{f\left(x', y', z', t \pm \frac{r}{v}\right)}{r}\, dV \tag{28.16†}$$

where $dV = dx'\, dy'\, dz'$ is an element of the volume of integration, and $r = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}$ is the distance between the point of integration and the point at which the value of Φ is calculated. The most significant fact in this formula is the different value of time in the functions Φ and f. The function f describes the source which creates the field described by Φ. First, we shall take the minus sign in the time argument of f, i.e., $t - r/v$. Equation (28.16) states that the value of Φ at the point x, y, z, at time t depends on the value of f at other points x', y', z' not at

† Let us consider the equation

$$\nabla^2\Phi - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\Phi = f(x, y, z, t) \tag{1}$$

and we wish to obtain a particular solution of the inhomogeneous equation. First of all, let us notice that if we have a solution of

$$\nabla^2 G - \frac{1}{v^2}\frac{\partial^2 G}{\partial t^2} = \delta(\mathbf{x} - \mathbf{x}')\,\delta(t - t') \tag{2}$$

where $\mathbf{x}'$, t' are fixed parameters, we have a solution to the original equation given by

$$\Phi = \int G(\mathbf{x}, t; \mathbf{x}'t')\, f(\mathbf{x}', t')\, dV'\, dt' \tag{3}$$

since

$$\int \delta(\mathbf{x} - \mathbf{x}')\, \delta(t - t')\, f(\mathbf{x}', t')\, dV'\, dt' = f(\mathbf{x}, t)$$

Thus, let us obtain a solution to (2). Notice that except at $\mathbf{x}'$ and t', G satisfies

$$\left(\nabla^2 - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right) G = 0$$

For the moment, let $\mathbf{x}' = 0$, $t' = 0$. Then, let us first look for solutions to

$$\frac{1}{r}\frac{\partial^2}{\partial r^2} rG - \frac{1}{v^2}\frac{\partial^2}{\partial t^2} G = 0$$

If $rG = g$, g satisfies

$$\left(\frac{\partial^2}{\partial r^2} - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right) g = 0$$

This is the same equation as (28.11) and thus

$$g(r, t) = a(t - r/v) + b(t + r/v)$$

Let us assume $g = g(t - r/v)$

Then

$$G = \frac{g(t - r/v)}{r}$$

$$\nabla^2 G = g\nabla^2\left(\frac{1}{r}\right) + 2\nabla\left(\frac{1}{r}\right)\cdot\nabla g + \frac{1}{r}\nabla^2 g$$

$$= 4\pi g\, \delta(\mathbf{x}) - \frac{2}{r^2}\frac{\partial g}{\partial r} + \frac{2}{r^2}\frac{\partial g}{\partial r} + \frac{1}{r}\frac{\partial^2 g}{\partial r^2}$$

$$= 4\pi g(t)\, \delta(\mathbf{x}) + \frac{1}{r}\frac{\partial^2 g}{\partial r^2}$$

Thus

$$\left(\nabla^2 - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right)\frac{g}{r} = 4\pi g(t)\, \delta(\mathbf{x}) + \frac{1}{r}\left(\frac{\partial^2}{\partial r^2} - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right) g(t - r/v)$$

$$= 4\pi g(t)\, \delta(\mathbf{x})$$

For $\frac{g}{r}$ to equal G

$$4\pi g(t)\delta(\mathbf{x}) = \delta(\mathbf{x})\, \delta(t)$$

or

$$g(t) = \frac{1}{4\pi}\delta(t)$$

Thus

$$G(r, t) = \frac{1}{4\pi}\frac{\delta(t - r/v)}{r}$$

and changing the singular point from (0, 0) to $(\mathbf{x}', t')$

$$G(\mathbf{x}, t; \mathbf{x}', t') = \frac{1}{4\pi}\frac{\delta\left(t - t' - \frac{|\mathbf{x} - \mathbf{x}'|}{v}\right)}{|\mathbf{x} - \mathbf{x}'|}$$

Hence, finally

$$\Phi = \frac{1}{4\pi}\int \frac{\delta\left(t - t' - \frac{|\mathbf{x} - \mathbf{x}'|}{v}\right)}{|\mathbf{x} - \mathbf{x}'|} f(\mathbf{x}', t')\, dV'\, dt'$$

$$= \frac{1}{4\pi}\int \frac{f\left(\mathbf{x}', t - \frac{|\mathbf{x} - \mathbf{x}'|}{v}\right)}{|\mathbf{x} - \mathbf{x}'|}\, dV'$$

the same time, but at an earlier time $t - r/v$. This time lag r/v is equal to the time taken by a signal, traveling with velocity v, to travel from x', y', z' to x, y, z. If we apply this to an electromagnetic field, then we can assert that the velocity v, given by equation (28.15), is the velocity of propagation of electromagnetic interactions, and that equation (28.16) takes into account the fact that this velocity is finite. The solution of (28.16) is called the *retarded solution*. The plus sign in the time argument of f, i.e., $t + r/v$, leads to a solution with the following meaning. The value of Φ at time t is determined by the values of f at other points of space at some later time, and not by the values at f at some earlier time. This is called the *advanced solution*, and has no clear physical meaning (unlike the retarded solution) and is used very rarely.

Retarded and Advanced Potentials. Since (28.16) is the solution of (28.9), the solutions of (28.7) and (28.8) may be written:

(1) in the form of retarded potentials

$$\mathbf{A}(x, y, z, t) = \frac{\mu}{4\pi} \int \frac{\mathbf{j}\left(x', y', z', t - \frac{r}{v}\right)}{r} dV \tag{28.17}$$

$$\varphi(x, y, z, t) = \frac{1}{4\pi\epsilon} \int \frac{\rho\left(x', y', z', t - \frac{r}{v}\right)}{r} dV \tag{28.18}$$

(2) in the form of advanced potentials

$$\mathbf{A}(x, y, z, t) = \frac{\mu}{4\pi} \int \frac{\mathbf{j}\left(x', y', z', t + \frac{r}{v}\right)}{r} dV \tag{28.19}$$

$$\varphi(x, y, z, t) = \frac{1}{4\pi\epsilon} \int \frac{\rho\left(x', y', z', t + \frac{r}{v}\right)}{r} dV \tag{28.20}$$

In these formulas the velocity V is given by (28.15).

The retarded potentials are of greater importance, and their physical meaning has been given above.

The existence of electromagnetic waves follows theoretically from the fact that the potentials of the electromagnetic field obey D'Alembert's equation, which may be solved in the form of waves.

§29. Radiation of a Linear Oscillator

The Hertz Oscillator. The simplest radiators of electromagnetic waves are the Hertz oscillator and a current-carrying loop. The Hertz oscillator

consists of two metal spheres joined by a conductor (Fig. 37). If there are equal and opposite charges on the spheres, and the system is left alone, an oscillatory charge exchange between the spheres will take place. A

Fig. 37

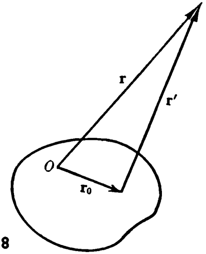


Fig. 38

current, which varies periodically, will flow in the conductor between the spheres. The oscillations of the current will be damped, but, if the resistance of the conductor is small, the damping may be ignored for a large number of periods. It follows that, at distances much greater than the distance between the spheres, the electromagnetic field of Hertz's oscillator may be described as the field of a dipole with a moment $\mathbf{p}$ varying with time.

Scalar Potential of a Dipole Varying with Time. We shall take the origin of coordinates at some point O close to the dipole (Fig. 38). The radius vector of the field point where the potential is to be measured will be denoted by $\mathbf{r}$. The remaining quantities are shown in Fig. 38. From equation (28.18) we may write

$$\varphi(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho\left(\mathbf{r}_0, t - \dfrac{r'}{c}\right)}{r'} \, dV \tag{29.1}$$

where $dV = dx_0\, dy_0\, dz_0$ and r' is the distance from the point of integration to the field point (Fig. 38). We shall consider the potential at large distances from the charged system $r_0/r \ll 1$.

Using the fact that

$$\mathbf{r}' = \mathbf{r} - \mathbf{r}_0 \qquad r' = \sqrt{r^2 - 2\mathbf{r}\cdot\mathbf{r}_0 + r_0^2}$$

we may expand the expression for r' as a series in powers of (r_0/r) and take only the linear term of the expansion

$$r' = r\left(1 - 2\frac{\mathbf{r}\cdot\mathbf{r}_0}{r^2} + \frac{r_0^2}{r^2}\right)^{1/2} = r - \frac{\mathbf{r}\cdot\mathbf{r}_0}{r} + \cdots \tag{29.2}$$

Using (29.2), we may expand the integrand in (29.1) as a Taylor series for r

$$\frac{\rho\left(\mathbf{r}_0, t - \frac{r'}{c}\right)}{r'} = \frac{\rho\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{r} - \frac{\mathbf{r}\cdot\mathbf{r}_0}{r}\frac{\partial}{\partial r}\left\{\frac{\rho\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{r}\right\}$$

$$+ \cdots = \frac{\rho}{r} - \frac{\mathbf{r}}{r}\cdot\frac{\partial}{\partial r}\left(\frac{\mathbf{r}_0\rho}{r}\right) + \cdots \tag{29.3}$$

Of course, we can leave out all terms in the expansion except the first only if these terms are small. The condition that the terms of the expansion are small will be used later to deduce the criterion of applicability of these formulas.

Substituting (29.3) in (29.1), we find

$$\varphi = \frac{1}{4\pi\epsilon_0 r}\int_V \rho\, dV - \frac{1}{4\pi\epsilon_0}\frac{\mathbf{r}}{r}\cdot\frac{\partial}{\partial r}\frac{1}{r}\int_V \mathbf{r}_0\rho\, dV \tag{29.4}$$

The integral in the first term on the right-hand side is equal to zero since the system is neutral, and the integral in the second term is the moment of the dipole

$$\int_V \mathbf{r}_0\rho\left(\mathbf{r}_0, t - \frac{r}{c}\right) dV = \mathbf{p}\left(t - \frac{r}{c}\right) \tag{29.5}$$

Consequently, the potential of the dipole takes the form

$$\varphi(\mathbf{r}, t) = -\frac{1}{4\pi\epsilon_0}\frac{\mathbf{r}}{r}\cdot\frac{\partial}{\partial r}\left\{\frac{\mathbf{p}\left(t - \frac{r}{c}\right)}{r}\right\} \tag{29.6}$$

This expression may be rewritten

$$\varphi(\mathbf{r}, t) = -\frac{1}{4\pi\epsilon_0}\operatorname{div}\frac{\mathbf{p}\left(t - \frac{r}{c}\right)}{r} \tag{29.7}$$

which may easily be verified by writing the divergence in spherical coordinates.

Vector Potential. On the basis of (28.17), we write

$$\mathbf{A}\ (\mathbf{r}, t) = \frac{\mu_0}{4\pi}\int_V \frac{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r'}{c}\right)}{r'} dV \tag{29.8}$$

and expand the integrand in series, as in (29.3)

$$\frac{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r'}{c}\right)}{r'} = \frac{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{r} - \frac{\mathbf{r}\cdot\mathbf{r}_0}{r}\frac{\partial}{\partial r}\left\{\frac{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{r}\right\} + \cdots$$

obtaining

$$\mathbf{A}\ (\mathbf{r}, t) = \frac{\mu_0}{4\pi r}\int_V \mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right) dV - \frac{\mu_0}{4\pi}\int_V \frac{\mathbf{r}\cdot\mathbf{r}_0}{r}\frac{\partial}{\partial r}\ \left(\frac{\mathbf{j}}{r}\right) dV + \cdots \tag{29.9}$$

In the case under consideration, currents are not closed, and hence, the first integral is not equal to zero. In fact, this integral makes the principal contribution to the potential. To evaluate this integral, we differentiate (29.5) with respect to time

$$\frac{\partial \mathbf{p}\left(t - \frac{r}{c}\right)}{\partial t} = \int \mathbf{r}_0\ \frac{\partial \rho\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{\partial t} dV$$

using the equation of continuity (4.6)

$$\frac{\partial \rho}{\partial t} = -\operatorname{div} \mathbf{j}$$

in which

$$\operatorname{div} \mathbf{j} = \frac{\partial j_x}{\partial x_0} + \frac{\partial j_y}{\partial y_0} + \frac{\partial j_z}{\partial z_0}$$

where x_0, y_0, and z_0 is the component of the radius vector $\mathbf{r}_0$. Consequently

$$\frac{\partial \mathbf{p}}{\partial t} = -\int_V \mathbf{r}_0 \operatorname{div} \mathbf{j}\, dV \tag{29.10}$$

Multiplying both sides of the equation by some arbitrary constant vector $\mathbf{a}$

$$\mathbf{a}\cdot\frac{\partial \mathbf{p}}{\partial t} = -\int_V (\mathbf{a}\cdot\mathbf{r}_0) \operatorname{div} \mathbf{j}\, dV \tag{29.11}$$

and transforming the integrand by the vector analysis formula (A.13) of Appendix 1, written in the form

$$(\mathbf{a}\cdot\mathbf{r}_0) \operatorname{div} \mathbf{j} = \operatorname{div} \mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0) - \mathbf{j}\cdot\operatorname{grad}(\mathbf{a}\cdot\mathbf{r}_0) = \operatorname{div} \mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0) - \mathbf{a}\cdot\mathbf{j}$$

we obtain

$$\mathbf{a}\cdot\frac{\partial \mathbf{p}}{\partial t} = -\int_V \operatorname{div} \mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0)\, dV + \mathbf{a}\cdot\int_V \mathbf{j}\, dV \tag{29.12}$$

The first integral on the right-hand side may be transformed in accordance with Gauss' theorem

$$\int_V \operatorname{div} \mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0)\, dV = \oint_S d\mathbf{S}\cdot\mathbf{j}\ (\mathbf{a}\cdot\mathbf{r}_0) = 0$$

where the surface S encloses the volume V under consideration. The integral over S equals zero, since all currents are within V and do not flow through S, i.e., $\mathbf{j} = 0$ at all points of S. Hence, instead of (29.12), we have

$$\mathbf{a}\cdot\frac{\partial \mathbf{p}}{\partial t} = \mathbf{a}\cdot\int_V \mathbf{j}\, dV \tag{29.13}$$

Since $\mathbf{a}$ is arbitrary, it follows that

$$\frac{\partial \mathbf{p}\left(t - \frac{r}{c}\right)}{\partial t} = \int_V \mathbf{j}\ \left(\mathbf{r}_0, t - \frac{r}{c}\right) dV \tag{29.14}$$

Then, from (29.9) and (29.14), we finally obtain the following expression for the vector potential

$$\mathbf{A}\ (\mathbf{r}, t) = \frac{\mu_0}{4\pi}\frac{\partial}{\partial t}\left\{\frac{\mathbf{p}\left(t - \frac{r}{c}\right)}{r}\right\} \tag{29.15}$$

Vector Potential of a Current-Carrying Loop. In a loop, the currents are closed and

$$\operatorname{div} \mathbf{j} = 0$$

From the equation of continuity (4.6) it follows that in this case the current density does not change with time. Therefore, the scalar potential is constant with respect to time, and is of no interest in the discussion of time-varying fields. The vector potential, however, is given by (29.8), from which (29.9) is obtained. In the case of a dipole varying with time, the currents are not closed, and the first term in (29.9) is the principal term, evaluation of which leads to (29.15). In the case of a current loop, however, the currents are closed; hence, the first integral in (29.9) becomes zero. Thus, the vector potential is given by the second term of this equation

$$\mathbf{A}\ (\mathbf{r}, t) = -\frac{\mu_0}{4\pi}\int_V \frac{\mathbf{r}\cdot\mathbf{r}_0}{r}\frac{\partial}{\partial r}\left(\frac{\mathbf{j}}{r}\right) dV \tag{29.16}$$

Since

$$\frac{\partial}{\partial r}\frac{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{r} = -\frac{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right)}{r^2} - \frac{1}{cr}\frac{\partial}{\partial t}\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right)$$

we may rewrite (29.6) in the form

$$\mathbf{A} = \frac{\mu_0}{4\pi}\int_V \frac{\mathbf{r}\cdot\mathbf{r}_0}{r^3}\left\{\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right) + \frac{r}{c}\frac{\partial}{\partial t}\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right)\right\} dV \tag{29.17}$$

In passing from (19.22) to (19.31) in §19, it was shown that

$$\int_V \frac{\mathbf{r}\cdot\mathbf{r}_0}{r^3}\mathbf{j}\left(\mathbf{r}_0, t - \frac{r}{c}\right) dV = \frac{\mathbf{M}\left(t - \frac{r}{c}\right)\times\mathbf{r}}{r^3} \tag{29.18}$$

where $\mathbf{M}$ is the magnetic moment of a closed current, defined by the equation (19.30). Using (29.18) to calculate (29.17), we obtain

$$\mathbf{A} = \frac{\mu_0}{4\pi}\frac{\mathbf{M}\left(t - \frac{r}{c}\right)\times\mathbf{r}}{r^3} + \frac{\mu_0}{4\pi}\frac{1}{r^2c}\frac{\partial\mathbf{M}\left(t - \frac{r}{c}\right)}{\partial t}\times\mathbf{r} \tag{29.19}$$

It must be pointed out that the field of the magnetic moment plays the principal role only if there is no radiation of the electric moment of the system. If the electric moment of the system generates radiation, then the electric moment plays the dominant role and the field of the time-varying magnetic moment may always be ignored. From a comparison of the formulas

$$\mathbf{M} = \frac{1}{2}\int \mathbf{r}_0 \times \rho\mathbf{v}\, dV = \frac{1}{2}\int \rho\mathbf{r}_0 \times \mathbf{v}\, dV$$

$$\mathbf{p} = \int \rho\mathbf{r}_0\, dV$$

it follows that

$$M \sim vp \tag{29.20}$$

where v is the velocity of motion of the charges. We now compare the values of the vector potential of a dipole $\mathbf{A}_d$ given by (29.15) and the vector potential $\mathbf{A}_M$ of a current loop given by (29.19). Taking only the orders of the quantities into account, we have

$$A_d \sim \frac{\mu_0}{4\pi}\frac{1}{r}\frac{\partial\mathbf{p}}{\partial t} \qquad A_M \sim \frac{\mu_0}{4\pi}\frac{1}{r\,c}\frac{\partial M}{\partial t}$$

Using (29.20), we see that

$$A_M \sim \frac{v}{c}A_d$$

Hence, for low velocities of motion of charges, $v \ll c$, the inequality

$$A_M \ll A_d$$

is always satisfied. Since the electric field intensity is defined as

$$E \sim \frac{\partial A}{\partial t}$$

we have, for $v \ll c$

$$E_M \ll E_d$$

i.e., the radiation of a magnetic moment may be ignored in comparison with the radiation of a dipole moment.

Electric and Magnetic Field of a Linear Oscillator. A *linear oscillator* is a dipole, the moment of which varies according to the law

$$\mathbf{p}(t) = \mathbf{p}_0 f(t)$$

where $\mathbf{p}_0$ is a constant vector and $f(t)$ is a periodic function. For convenience, in the subsequent calculations, we introduce the vector

$$\mathbf{\Pi}(t, r) \equiv \frac{\mathbf{p}_0 f\left(t - \frac{r}{c}\right)}{r} \equiv \mathbf{p}_0 \Phi(t, r) \tag{29.21}$$

which is called the *Hertz vector*, or the *polarization potential*. It satisfies the equation

$$\Delta \mathbf{\Pi} - \frac{1}{c^2} \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} = 0 \tag{29.22}$$

From equations (29.7) and (29.15) for the scalar and vector potentials, we obtain

$$\mathbf{B} = \operatorname{curl} \mathbf{A} = \frac{\mu_0}{4\pi} \operatorname{curl} \frac{\partial \mathbf{\Pi}}{\partial t} = \frac{\mu_0}{4\pi} \frac{\partial}{\partial t} \operatorname{curl} \mathbf{\Pi} \tag{29.23}$$

$$\begin{aligned} \mathbf{E} &= -\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} = \frac{1}{4\pi\epsilon_0} \operatorname{grad} \operatorname{div} \mathbf{\Pi} - \frac{\mu_0}{4\pi} \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} \\ &= \frac{1}{4\pi\epsilon_0} \left(\operatorname{grad} \operatorname{div} \mathbf{\Pi} - \frac{1}{c^2} \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} \right) = \frac{1}{4\pi\epsilon_0} \operatorname{curl} \operatorname{curl} \mathbf{\Pi} \end{aligned} \tag{29.24}$$

using (29.22) and the vector analysis formula (A.8) of Appendix 1, and remembering that $\mu_0 \epsilon_0 = 1/c^2$.

The value of $\operatorname{curl} \mathbf{\Pi}$ is found by using formula (A.14) of Appendix 1

$$\operatorname{curl} \mathbf{\Pi} = \operatorname{curl} \mathbf{p}_0 \Phi = \operatorname{grad} \Phi \times \mathbf{p}_0 = \frac{1}{r} \frac{\partial \Phi}{\partial r} \mathbf{r} \times \mathbf{p}_0 \tag{29.25}$$

Further calculations may be performed more easily in spherical coordinates. We shall take the polar axis z of the spherical system of coordinates along the vector $\mathbf{p}_0$, placing the origin at the center of the dipole. The polar and azimuthal angles are denoted, respectively, by Θ and α (Fig. 39). Clearly

$$(\mathbf{r} \times \mathbf{p}_0)_r = (\mathbf{r} \times \mathbf{p}_0)_\Theta = 0 \qquad (\mathbf{r} \times \mathbf{p}_0)_\alpha = -r p_0 \sin \Theta$$

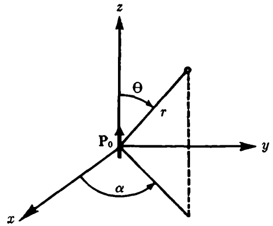


Fig. 39

Therefore, we obtain

$$\operatorname{curl}_r \mathbf{\Pi} = \operatorname{curl}_\Theta \mathbf{\Pi} = 0 \qquad \operatorname{curl}_\alpha \mathbf{\Pi} = -\sin\Theta \frac{\partial \Pi}{\partial r}$$

from which, on the basis of equation (29.23), we find the following expressions for the components of the magnetic field in spherical coordinates

$$B_r = B_\Theta = 0 \qquad B_\alpha = \frac{\mu_0}{4\pi}\frac{\partial}{\partial t}\operatorname{curl}_\alpha \mathbf{\Pi} = -\frac{\mu_0}{4\pi}\sin\Theta\frac{\partial^2 \Pi}{\partial t\,\partial r} \tag{29.26}$$

The components of the electric field are calculated using the formula for curl in spherical coordinates

$$\left.\begin{aligned} E_r &= \frac{1}{4\pi\epsilon_0}\frac{1}{r\sin\Theta}\frac{\partial}{\partial\Theta}(\sin\Theta\operatorname{curl}_\alpha \mathbf{\Pi}) = -\frac{1}{2\pi\epsilon_0}\frac{\cos\Theta}{r}\frac{\partial\Pi}{\partial r} \\ E_\Theta &= -\frac{1}{4\pi\epsilon_0}\frac{1}{r}\frac{\partial}{\partial r}(r\operatorname{curl}_\alpha \mathbf{\Pi}) = \frac{1}{4\pi\epsilon_0}\frac{\sin\Theta}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Pi}{\partial r}\right) \\ E_\alpha &= 0 \end{aligned}\right\} \tag{29.27}$$

Equations (29.26) and (29.27) show that the electric field vector lies in meridional planes, and the magnetic field vector is perpendicular to meridional planes drawn through the corresponding points, while the magnetic lines of force coincide with parallels of the spherical system. The electric and magnetic fields are perpendicular to each other at every point.

Equations (29.26) and (29.27) hold for an arbitrary law of variation of the dipole moment with time. If the dipole moment varies according to the harmonic law

$$\mathbf{p} = \mathbf{p}_0 e^{i\omega t} \tag{29.27a}$$

then the Hertz vector may be written

$$\mathbf{\Pi} = \mathbf{p}_0 \frac{e^{i\omega\left(t-\frac{r}{c}\right)}}{r} \tag{29.28}$$

Carrying out the differentiations in equations (29.26) and (29.27), we find the following expressions for the components of the electric and magnetic fields that are different from zero

$$\left.\begin{aligned} B_\alpha &= \frac{\mu_0}{4\pi} i\omega \sin\Theta \left(\frac{1}{r} + \frac{i\omega}{c}\right) \Pi \\ E_r &= \frac{1}{2\pi\epsilon_0} \cos\Theta \left(\frac{1}{r^2} + \frac{i\omega}{cr}\right) \Pi \\ E_\Theta &= \frac{1}{4\pi\epsilon_0} \sin\Theta \left(\frac{1}{r^2} + \frac{i\omega}{cr} - \frac{\omega^2}{c^2}\right) \Pi \end{aligned}\right\} \tag{29.29}$$

The field in the immediate proximity of the oscillator at distances less than a wavelength $\lambda = cT = 2\pi c/\omega$ is the same as the field of a static dipole and a current. The field is more interesting in regions of space at distances many times greater than the wavelength of the radiated wave, i.e., when $r \gg \lambda$. This region is called the *wave zone*, or the *radiative field.*

Field of an Oscillator in the "Wave Zone." The distance r of a point in the wave zone, by definition, satisfies

$$\frac{1}{r} \ll \frac{\omega}{c} \tag{29.30}$$

Hence, in equations (29.29) we may ignore $1/r$ and $1/r^2$ in comparison with ω/c and ω^2/c^2. We then obtain the following expressions for the components of $\mathbf{E}$ and $\mathbf{B}$

$$\left.\begin{aligned} B_\alpha &= -\frac{\mu_0}{4\pi}\frac{\omega^2}{c} \sin\Theta\, \Pi \qquad & B_r &= B_\Theta = 0 \\ E_\Theta &= -\frac{1}{4\pi\epsilon_0}\frac{\omega^2}{c^2} \sin\Theta\, \Pi \qquad & E_r &= E_\alpha = 0 \end{aligned}\right\} \tag{29.31}$$

where we may take Π to be either the real part or the imaginary part of (29.28), e.g.

$$\Pi = \frac{p_0 \cos\omega\left(t - \frac{r}{c}\right)}{r} \tag{29.32}$$

Thus, the electromagnetic radiation field of an oscillator in the wave zone may be written

$$\left.\begin{aligned} E_\Theta &= cB_\alpha = -\frac{1}{4\pi\epsilon_0}\frac{\omega^2}{c^2}\frac{\sin\Theta}{r}p_0\cos\omega\left(t-\frac{r}{c}\right) \\ E_r &= E_\alpha = 0 \qquad B_r = B_\Theta = 0 \end{aligned}\right\} \tag{29.33}$$

From (29.33) it follows that

$$\sqrt{\epsilon_0}E_\Theta = \sqrt{\mu_0}H_\alpha \tag{29.34}$$

In the Gaussian absolute system of units, this expression has the form of an equality of the electric field and magnetic field vectors. In the SI international system, these vectors have different dimensions, and it is meaningless to compare their values. Equation (29.33) states that in the wave zone, the electric and magnetic fields are perpendicular to one another and to the radius vector $\mathbf{r}$. The vectors $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{r}$ form a right-hand screw triad. The magnitudes of $\mathbf{E}$ and $\mathbf{B}$ decrease in inverse proportion to the first power of the distance. The wave represented by (29.33) is called a *spherical wave*. The constant phase surfaces of this wave are spheres. The phase velocity of the wave is equal to the velocity of light c. Thus, the oscillator radiates a spherical electromagnetic wave. At large distances from the oscillator, small sections of the wave may be considered to be plane electromagnetic waves. The properties of these plane waves will be considered in greater detail in the following chapter.

Energy Radiated by an Oscillator. The flow of the electromagnetic energy is described by the Poynting vector

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \tag{29.35}$$

Let us evaluate the energy flux Q through the surface S of a sphere of radius r. We have

$$\begin{aligned} Q = \oint_S \mathbf{E}\times\mathbf{H}\cdot d\mathbf{S} &= \int_S E_\Theta H_\alpha\, dS \\ &= \frac{1}{16\pi^2\epsilon_0}\frac{\omega^4 p_0^2}{c^3}\cos^2\omega\left(t-\frac{r}{c}\right)\int_0^\pi \sin^3\Theta\, d\Theta \int_0^{2\pi} d\alpha \\ &= \frac{1}{6\pi\epsilon_0}\frac{\omega^4 p_0^2}{c^3}\cos^2\omega\left(t-\frac{r}{c}\right) \end{aligned} \tag{29.36}$$

This is the rate of energy flow through the surface of the sphere, i.e., the energy flowing per unit time. This energy is radiated by the oscillator. The mean rate of energy flow from the oscillator in one period is equal to

$$\langle Q\rangle = \frac{1}{T}\int_0^T Q\, dt = \frac{1}{12\pi\epsilon_0}\frac{\omega^4 p_0^2}{c^3} \tag{29.37}$$

Thus, the power radiated by the oscillator depends very strongly on the frequency, and is proportional to the fourth power of the frequency. This

means that to increase the radiated power, it is advantageous to use short waves.

Since the Poynting vector is inversely proportional to the square of the distance, while the surface of the sphere is directly proportional to the square of the distance, the value of the energy flux crossing the sphere does not vary with distance. Thus, the energy travels without any loss from the oscillator to distant regions of space in the form of energy of electromagnetic waves radiated by the oscillator. Owing to the loss of the energy radiated, the oscillations of the oscillator must be damped. To maintain undamped oscillations, it is necessary to have a constant external supply of energy.

§30. Radiation of a Current Loop

Since a current loop has no dipole moment, and, consequently, the scalar potential is equal to zero, the field depends on only the vector potential, and is given by the formulas

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} \tag{30.1}$$

$$\mathbf{B} = \operatorname{curl} \mathbf{A} \tag{30.2}$$

The expression for the vector potential of a current loop is given by (29.19). There are two terms in this formula, the first of which decreases with distance as $1/r^2$, and the second as $1/r$. In the wave zone, the field described by the first term is negligibly small in comparison with the field described by the second term. Hence, when one considers radiation, he need take only the second term into account.

The electric field intensity is, according to (30.1), given by

$$\mathbf{E} = -\frac{\mu_0}{4\pi}\frac{1}{cr^2}\frac{\partial^2 \mathbf{M}\left(t - \frac{r}{c}\right)}{\partial^2 t} \times \mathbf{r} \tag{30.3}$$

We shall take the magnetic moment of the current loop to be

$$\mathbf{M}(t) = \mathbf{M}_0 \cos \omega t \tag{30.4}$$

Ignoring the first term, the vector potential (29.19) becomes

$$\mathbf{A} = -\frac{\mu_0}{4\pi}\frac{\omega}{c}\frac{\sin \omega\left(t - \frac{r}{c}\right)}{r^2}(\mathbf{M}_0 \times \mathbf{r}) \tag{30.5}$$

In spherical coordinates (Fig. 39), the components of the vector potential (30.5) are equal to

$$A_r = A_\Theta = 0$$

$$A_\alpha = -\frac{\mu_0}{4\pi}\frac{\omega}{c} M_0 \sin\Theta \frac{\sin\omega\left(t - \frac{r}{c}\right)}{r}$$

Using the expression for curl in spherical coordinates and dropping terms of the order of $1/r^2$, we obtain

$$\left.\begin{aligned} B_r &= 0 \\ B_\Theta = -\frac{1}{r}\frac{\partial}{\partial r}(r A_\alpha) &= -\frac{\mu_0}{4\pi}\frac{\omega^2}{c^2} M_0 \sin\Theta \frac{\cos\omega\left(t - \frac{r}{c}\right)}{r} \\ B_\alpha &= 0 \end{aligned}\right\} \tag{30.6}$$

According to (30.3) and (30.4), the components of **E** in spherical coordinates are

$$E_r = E_\Theta = 0$$

$$E_\alpha = \frac{\mu_0}{4\pi}\frac{\omega^2}{c} M_0 \frac{\sin\Theta}{r} \cos\omega\left(t - \frac{r}{c}\right) \tag{30.7}$$

Comparison of (30.7) and (30.6) shows that for the nonzero components of the field the following relationship holds

$$E_\alpha = -cB_\Theta = \frac{\mu_0}{4\pi}\frac{\omega^2}{c}\frac{\sin\Theta}{r} M_0 \cos\omega\left(t - \frac{r}{c}\right) \tag{30.8}$$

Comparing this expression with equation (29.33) for a dipole, we see that if

$$M_0 = cp_0 \tag{30.9}$$

then the electric field and the magnetic induction of a dipole are equal to the electric field and the magnetic induction of a loop. Only the directions of the vectors are different. For a dipole the electric field is directed along the meridians, and, for a loop, this field is perpendicular to the meridional planes. The orientation of the magnetic induction differs in a similar manner. As is evident from a comparison of (30.8) and (29.33), the electromagnetic field vectors of a dipole and a loop are related by

$$\begin{aligned} E_\alpha\,(\text{loop}) &= -cB_\alpha\,(\text{dipole}) \\ cB_\Theta\,(\text{loop}) &= E_\Theta\,(\text{dipole}) \end{aligned}$$

The power radiated by a current loop is given by (29.36) and (29.37) where the dipole moment is replaced by the magnetic moment, in accordance with equation (30.9).

§31. Directed Radiation

The flow of radiation energy is given by the Poynting vector

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \tag{31.1}$$

As equations (29.33) and (30.8) show, the radiation of the simplest radiators is directional, since

$$|\mathbf{P}| \sim \sin^2 \Theta \tag{31.2}$$

Thus, the radiation is maximum in the direction perpendicular to the dipole moment or the current loop. There is no radiation in the direction of the moment.

However, this property of the radiation is not very marked in elementary radiators which have been considered thus far. It may be made stronger, however, if we use certain radiators in which the phases of the oscillations are displaced with respect to one another in some definite way. Interference occurs between the waves radiated by different oscillators, so that along some directions the waves are strengthened, and along others they are weakened. In principle, the effect is the same as the well-known interference of light that has passed through a diffraction grating; the beams are strengthened along some directions and weakened along others. Using such interference, it is possible to generate very strongly directed beams of electromagnetic waves.

Without going into the technical details of how this may be produced, we note that the devices used for producing such radiation have dimensions of the order of one wavelength. This follows from the simple fact that the distance between elementary radiators must be of the order of one wavelength for interference to occur. Hence, in order to keep the size of the generating apparatus fairly small, we must use sufficiently short electromagnetic waves.

At present, the technique of obtaining directed radiation has reached a very high level of development and is widely used (radar devices, radiotelescopes). Very recently, devices have been developed which give practically parallel beams of light (lasers). These are based on the so-called stimulated radiation of atomic systems. This goes beyond the scope of the present volume, but can be found in treatises on quantum mechanics.

PROBLEMS

1 Determine the mean total radiation flux of a current loop, $I = I_0 \cos \omega t$. The area of the loop is S.

Consider the numerical example: $I_0 = 10$ amp, $S = 100\ \text{cm}^2$, $\omega = 10^8\ \text{sec}^{-1}$.

Answer: $\langle Q\rangle = \dfrac{\mu_0}{12\pi c^3}\omega^4 I_0{}^2 S^2 = 0.124\ W$

2 Find the maximum value of the radiation flux density in the plane of the loop of example 1, at a distance 200 m from it.

Answer: $P = \dfrac{\mu_0}{16\pi^2 c^3}\dfrac{\omega^4 I_0{}^2 S^2}{r^2} = 0.47\cdot 10^{-5}\ W/\mathrm{m}^2 = 0.047\ W/\mathrm{cm}^2$

3 Determine the length of a dipole l whose radiated power is equal to the power radiated by the loop of example 1. The oscillation frequency of the dipole is equal to the oscillation frequency of the current in the loop, and the magnitude of each of the dipole charges is $q = 10^{-4}$ coul.

Answer: $l = \dfrac{I_0 S}{qc} = 3.3 \times 10^{-4}\ \mathrm{m} = 0.33\ \mathrm{mm}$

CHAPTER 6

Propagation of Electromagnetic Waves

§32. Propagation of Electromagnetic Waves in Dielectrics

Plane Monochromatic Waves. An electromagnetic wave is said to be *plane* if $\mathbf{E}$ and $\mathbf{H}$ have the same value at all points of any plane perpendicular to the direction of propagation of the waves. Thus, the surfaces of constant phase in a plane wave are planes perpendicular to the direction of propagation of the wave. A wave is said to be *monochromatic* if $\mathbf{E}$ and $\mathbf{H}$ vary with time according to the harmonic law and with a definite frequency.

For example, if a plane electromagnetic wave is propagated along the z axis, then $\mathbf{E}$ and $\mathbf{H}$ take the form

$$\mathbf{E}(z, t) = \mathbf{E}(z)e^{i\omega t} \qquad \mathbf{H}(z, t) = \mathbf{H}(z)e^{i\omega t} \tag{32.1}$$

Equations for E and H. Let us consider the case of a homogeneous unbounded medium $\epsilon = \text{const}$, $\mu = \text{const}$ in the absence of charges. The conductivity is $\lambda = 0$. We start from Maxwell's equations

$$\text{curl}\, \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t} \tag{32.2}$$

$$\text{curl}\, \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{32.3}$$

Differentiating both sides of (32.2) with respect to time and eliminating $\partial \mathbf{H}/\partial t$ from the left-hand side using (32.3), we find

$$-\frac{1}{\mu}\, \text{curl curl}\, \mathbf{E} = \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{32.4}$$

Using formula (A.8) of Appendix 1 and remembering that the divergence

of the electric field in a homogeneous medium in the absence of charges is equal to zero, we obtain, instead of (32.4), the equation

$$\Delta \mathbf{E} - \epsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \tag{32.5}$$

The equation for **H** is obtained in a similar manner, and, due to the symmetry of (32.2) and (32.3), this has the same form as equation (32.5)

$$\Delta \mathbf{H} - \epsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} = 0 \tag{32.6}$$

Thus, **E** and **H** satisfy the same wave equation with the same velocity of propagation

$$v = \frac{1}{\sqrt{\epsilon\mu}} = \frac{c}{\sqrt{\epsilon'\mu'}} \tag{32.7}$$

where ϵ' and μ' are the permittivity and permeability of the medium, respectively, and c is the velocity of light *in vacuo.*

Solution in the Form of Plane Monochromatic Waves. Let us take the z axis along the direction of propagation of an electromagnetic wave. In this case, **E** and **H** are given by (32.1). Let us consider, for example, the equation for **E**. We shall substitute equation (32.1) for **E** in equation (32.5). Differentiating with respect to time, and eliminating the exponential time factor, we obtain

$$\frac{d^2\mathbf{E}(z)}{dz^2} + k^2\mathbf{E}(z) = 0 \qquad k = \omega\sqrt{\epsilon\mu} \tag{32.8}$$

The general solution of this equation has the form

$$\mathbf{E}(z) = \mathbf{a}_1 e^{-ikz} + \mathbf{a}_2 e^{ikz} \tag{32.9}$$

Substituting (32.9) in (32.1), we find

$$\mathbf{E}(z, t) = \mathbf{a}_1 e^{i(\omega t - kz)} + \mathbf{a}_2 e^{i(\omega t + kz)} \tag{32.10}$$

The first term on the right-hand side of (32.10) represents a wave traveling in the positive direction of the z axis. This follows from the fact that the point of constant phase

$$\omega t - kz = \text{const} \tag{32.11}$$

moves in the direction of increasing values of z, i.e., as t increases the value of z in (32.11) also increases. Similar considerations show that the second term of (32.10) describes a wave traveling in the negative direction of the z axis.

The solution of (32.6) is found in a similar manner. Hence, we may write down the following expression for **E** and **H** for an electromagnetic wave traveling in the positive direction of the z axis

$$\mathbf{E}(z, t) = \mathbf{E}_0 e^{i(\omega t - kz)} \qquad \mathbf{H}(z, t) = \mathbf{H}_0 e^{i(\omega t - kz)} \tag{32.12}$$

Here, $\mathbf{E}_0$ and $\mathbf{H}_0$ are the amplitudes of the fields.

Equations (32.12) show that plane waves in a homogeneous dielectric are propagated without a change in amplitude, i.e., without damping. The phase velocity of the waves is found by means of the differentiating equation (32.11) with respect to time. It is equal to

$$\frac{dz}{dt} = \frac{\omega}{k} = \frac{1}{\sqrt{\epsilon\mu}} = \frac{c}{\sqrt{\epsilon'\mu'}} = v \tag{32.13}$$

The wavelength l is related to k by the equation

$$k = \omega\sqrt{\epsilon\mu} = \frac{\omega}{v} = \frac{2\pi}{l} \tag{32.14}$$

Equations (32.12) are formulated for a special choice of coordinates, with the z axis along the direction of propagation of the wave. To eliminate this restriction, we shall introduce the wave vector $\mathbf{k}$, whose direction lies in the direction of propagation of the wave, and whose magnitude is given by (32.8). From the definition of a plane wave traveling along the z axis, the magnitudes of E and H are equal at any point of a plane perpendicular to the z axis. Let the radius vector of some point on such a constant phase plane be $\mathbf{r}$. Clearly, $kz = \mathbf{k}\cdot\mathbf{r}$, and, therefore, instead of (32.12) we may write

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})} \qquad \mathbf{H}(\mathbf{r}, t) = \mathbf{H}_0 e^{i(\omega t - \mathbf{kr})} \tag{32.15}$$

These equations describe the electric and magnetic fields of a plane electromagnetic wave traveling in the direction of $\mathbf{k}$. The frequency of the wave is ω, and the wavelength l is given by (32.14).

To investigate the properties of plane waves, we shall substitute equation (32.15) in one of Maxwell's equations. Using the vector analysis operator **nabla** ∇, defined by

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{32.16}$$

the divergence and curl operations, applied to a vector $\mathbf{A}$, are written, respectively, as scalar and vector products of ∇ and $\mathbf{A}$

$$\operatorname{div}\mathbf{A} = \nabla\cdot\mathbf{A}, \ \operatorname{curl}\mathbf{A} = \nabla\times\mathbf{A} \tag{32.17}$$

The above equations may easily be verified using (32.16). Direct calculation shows that

$$\nabla e^{-i\mathbf{k}\cdot\mathbf{r}} = -i\mathbf{k}e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{32.18}$$

In the case of a homogeneous dielectric, there are no volume charges, and,

consequently, div E = 0. Substituting equation (32.15) for **E**, and using (32.17) and (32.18), we obtain

$$\operatorname{div} \mathbf{E} = \nabla\cdot\mathbf{E} = -i\mathbf{k}\cdot\mathbf{E} = 0$$

Similarly, from Maxwell's equation for a homogeneous medium, div **H** = 0, we have

$$\operatorname{div} \mathbf{H} = \nabla\cdot\mathbf{H} = -i\mathbf{k}\cdot\mathbf{H} = 0$$

From the scalar products

$$\mathbf{k}\cdot\mathbf{E} = 0 \qquad \mathbf{k}\cdot\mathbf{H} = 0 \tag{32.19}$$

it follows that in the case of a plane wave, **E** and **H** lie in a plane perpendicular to the direction of propagation of the wave.

If we substitute the values of **E** and **H** from (32.15) in Maxwell's equation (32.3), we obtain

$$-i\mathbf{k} \times \mathbf{E} = -i\omega\mu\mathbf{H} \tag{32.20}$$

Let **n** be a unit vector in the direction of propagation of the wave. Then, on the basis of (32.8), we may write

$$\mathbf{k} = k\mathbf{n} = \omega\sqrt{\epsilon\mu}\,\mathbf{n}$$

Substituting this expression for **k** in (32.20), we obtain the relationship

$$\sqrt{\epsilon}\,\mathbf{n} \times \mathbf{E} = \sqrt{\mu}\,\mathbf{H} \tag{32.21}$$

Hence, it is clear that **E** and **H** are perpendicular to each other. It has been shown that they are both perpendicular to **n**. From (32.21) it follows that **E**, **H**, and **n** make a right-hand screw triad of mutually perpendicular vectors (Fig. 40).

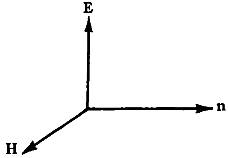


Fig. 40

Taking moduli of both sides of (32.21), we obtain

$$\sqrt{\epsilon}|\mathbf{E}| = \sqrt{\mu}|\mathbf{H}| \tag{32.22}$$

From (32.21), we may conclude that, in a homogeneous dielectric **E** and **H** in a plane wave are in phase.

The flux density of an electromagnetic wave is given by the Poynting vector, whose absolute magnitude for a plane wave is

$$|\mathbf{P}| = |\mathbf{E} \times \mathbf{H}| = |\mathbf{E}| \cdot |\mathbf{H}| = \frac{1}{\sqrt{\epsilon\mu}} \frac{1}{2} (\epsilon E^2 + \mu H^2)$$

where (32.22) has been used. Since $v = 1/\sqrt{\epsilon\mu}$ is the phase velocity of a plane wave, and

$$\frac{1}{2} (\epsilon E^2 + \mu H^2) = u$$

is the energy density of the electromagnetic field, we may write the expression for the Poynting vector

$$\mathbf{P} = u\mathbf{v} \tag{32.23}$$

Thus, the velocity of motion of the energy carried by a plane wave in a homogeneous dielectric is equal to the phase velocity of the wave. According to (32.15), a plane wave travels in a homogeneous dielectric without change of amplitude, i.e., without loss of energy.

§33. Propagation of Electromagnetic Waves in Conducting Media

Let us consider the case of a homogeneous unbounded conducting medium λ = const, μ = const, ϵ = const, $\lambda \neq 0$. Substituting in Maxwell's equations

$$\text{curl}\,\mathbf{H} = \mathbf{j} + \epsilon \frac{\partial \mathbf{E}}{\partial t} = \lambda \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \tag{33.1}$$

$$\text{curl}\,\mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{33.2}$$

the expressions for **E** and **H** from (32.15), we obtain

$$-i\mathbf{k} \times \mathbf{H} = i\omega \left(\epsilon + \frac{\lambda}{i\omega}\right) \mathbf{E} \tag{33.3}$$

$$-i\mathbf{k} \times \mathbf{E} = -i\omega\mu \mathbf{H} \tag{33.4}$$

If we put $\lambda = 0$ in (33.3), then we obtain the corresponding equation for a dielectric. Equation (33.4) does not differ from the corresponding equation for a dielectric. Thus, mathematically, the case of a conducting medium differs from that of a dielectric only in the replacement of the permittivity ϵ with the complex permittivity

$$\epsilon_\omega = \epsilon + \frac{\lambda}{i\omega} = \epsilon - i\frac{\lambda}{\omega} \tag{33.5}$$

All subsequent calculations are completely identical in form with the calculations for plane waves in a dielectric, provided that we replace ϵ

by ϵ_ω. Thus, instead of the real quantity k, we have the complex quantity k_ω, where

$$k_\omega{}^2 = \omega^2\epsilon_\omega\mu = \omega^2\epsilon\mu - i\omega\lambda\mu \tag{33.6}$$

Writing k_ω as a complex number

$$k_\omega = k - is \tag{33.7}$$

equation (33.6) may be rewritten

$$k^2 - 2iks - s^2 = \omega^2\epsilon\mu - i\omega\lambda\mu$$

Equating the real and imaginary parts of this equation, we find

$$k^2 - s^2 = \omega^2\epsilon\mu \equiv a \tag{33.8}$$

$$2ks = \omega\lambda\mu \equiv b \tag{33.9}$$

The solution of this algebraic system of equations is of the form

$$k^2 = \frac{a}{2}\left(\sqrt{1 + \frac{b^2}{a^2}} + 1\right) = \frac{\omega^2\epsilon\mu}{2}\left(\sqrt{1 + \left(\frac{\lambda}{\epsilon\omega}\right)^2} + 1\right) \tag{33.10}$$

$$s^2 = \frac{a}{2}\left(\sqrt{1 + \frac{b^2}{a^2}} - 1\right) = \frac{\omega^2\epsilon\mu}{2}\left(\sqrt{1 + \left(\frac{\lambda}{\epsilon\omega}\right)^2} - 1\right) \tag{33.11}$$

By analogy with equation (32.15), a solution in the form of a plane wave traveling in the conducting medium in the positive direction of the z axis, is written in the following form

$$\mathbf{E} = \mathbf{E}_0 e^{i(\omega t - k_\omega z)} = \mathbf{E}_0 e^{-sz}\, e^{i(\omega t - kz)} \tag{33.12}$$

$$\mathbf{H} = \mathbf{H}_0 e^{i(\omega t - k_\omega z)} = \mathbf{H}_0 e^{-sz}\, e^{i(\omega t - kz)} \tag{33.13}$$

The amplitude of this wave decreases. Consequently, in a conducting medium, the electromagnetic waves are damped. The amplitude decreases by a factor e over a path of length

$$\Delta = \frac{1}{s} \tag{33.14}$$

The quantity Δ describes the depth of penetration of the wave into the conducting medium.

Let us estimate the depth of penetration for different wavelengths. For visible light, the wavelengths are

$$l = (0.4 - 0.75)\, 10^{-4}\ \text{cm}$$

corresponding to frequencies of the order of $\omega = 5 \times 10^{15}\ \text{sec}^{-1}$. The conductivity of metals is $\lambda \approx 10^7\ \text{ohm}^{-1}\ \text{m}^{-1}$, and ϵ may be taken as approximately equal to ϵ_0. Thus, in this case

$$\frac{\lambda}{\epsilon\omega} \approx 2 \times 10^2 \gg 1$$

At lower frequencies, i.e., for longer wavelengths, this inequality is obeyed more strongly. Hence, in equation (33.11) we may neglect unity compared with $\lambda/\epsilon\omega$. Then, s becomes

$$s = \sqrt{\frac{1}{2}\,\omega\lambda\mu}$$

Hence, the depth of penetration equals

$$\Delta = \sqrt{\frac{2}{\omega\lambda\mu}} \tag{33.15}$$

Since the wavelength l is related to the frequency ω by the relationship $\omega = 2\pi/l\sqrt{\epsilon\mu}$, equation (33.15) may be rewritten as follows

$$\Delta = \sqrt{\frac{l}{\pi\lambda}}\left(\frac{\epsilon}{\mu}\right)^{1/4} \tag{33.16}$$

The quantity $\sqrt{\mu/\epsilon}$ has the dimensions of resistance, and is called the *wave impedance* of the medium. The wave impedance of empty space is equal to

$$\sqrt{\frac{\mu_0}{\epsilon_0}} = 377\ \Omega$$

As an example, let us consider copper, for which $\lambda = 5 \times 10^7\ \text{ohm}^{-1}\ \text{m}^{-1}$, $\mu \approx \mu_0$, $\epsilon \approx \epsilon_0$. For wavelength $l = 10^{-4}$ cm, the depth of penetration is equal to

$$\Delta \approx 4 \times 10^{-9}\ \text{m} = 4 \times 10^{-7}\ \text{cm}$$

Thus, in this case, the depth of penetration is many times less than the wavelength. Hence, strictly speaking, it is meaningless to talk of the propagation of an electromagnetic wave in a conducting medium if the wave is very rapidly damped. As is clear from equation (33.16), this conclusion remains true even for very long wavelengths, since the depth of penetration increases as the square root of the wavelength, i.e., more slowly than the wavelength.

The physical reason for this rapid damping in a conducting medium is the transformation of electromagnetic energy into Joule heat: the electric field of the waves generates conduction currents, and heat is liberated in accordance with the Joule-Lenz law.

§34. Refraction and Reflection of Plane Electromagnetic Waves at a Boundary Between Dielectrics

Boundary Conditions for the Electromagnetic Wave Vectors. The problem of refraction and reflection of plane waves at a boundary between dielectrics is solved by means of boundary conditions.

Consider two media (Fig. 41) separated by a plane surface. An electromagnetic wave is incident on B, the boundary from the side of medium 1, part of this wave being reflected into medium 1, and part being refracted

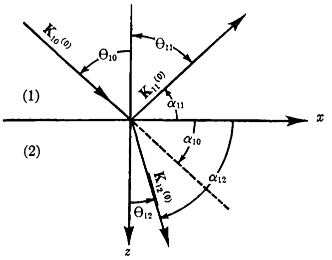


Fig. 41

and transmitted into medium 2. Thus, in medium 1, there is an incident and a reflected wave, and, in medium 2, a refracted wave.

We shall denote quantities referring to the incident wave by the suffix 10, to the reflected wave by 11, and to the refracted wave by 12. Then, the electric fields for the three waves are

$$\mathbf{E}_{10}(\mathbf{r}, t) = \mathbf{E}_{10}{}^{(0)} e^{i(\omega_{10}t - \mathbf{k}_{10}\cdot\mathbf{r})} \tag{34.1}$$

$$\mathbf{E}_{11}(\mathbf{r}, t) = \mathbf{E}_{11}{}^{(0)} e^{i(\omega_{11}t - \mathbf{k}_{11}\cdot\mathbf{r})} \tag{34.2}$$

$$\mathbf{E}_{12}(\mathbf{r}, t) = \mathbf{E}_{12}{}^{(0)} e^{i(\omega_{12}t - \mathbf{k}_{12}\cdot\mathbf{r})} \tag{34.3}$$

The expressions for the magnetic field are similar.

The boundary condition expressing the continuity of the tangential components of the electric field vector becomes

$$E_{10t}{}^{(0)} e^{i(\omega_{10}t - \mathbf{k}_{10}\cdot\mathbf{r})} + E_{11t}{}^{(0)} e^{i(\omega_{11}t - \mathbf{k}_{11}\cdot\mathbf{r})} = E_{12t}{}^{(0)} e^{i(\omega_{12}t - \mathbf{k}_{12}\cdot\mathbf{r})} \tag{34.4}$$

Frequency Invariance under Reflection and Refraction. For simplicity, we shall write (34.4) in the form

$$a e^{i\omega_{10}t} + b e^{i\omega_{11}t} = c e^{i\omega_{12}t} \tag{34.5}$$

where a, b, c are independent of time. Differentiating both sides of the equation with respect to t, we have

$$i\omega_{10}\, a e^{i\omega_{10}t} + i\omega_{11}\, b e^{i\omega_{11}t} = i\omega_{12}\, c e^{i\omega_{12}t} \tag{34.6}$$

Substituting the quantity $ce^{i\omega_{12}t}$ from (34.5) in (34.6), we obtain

$$ia(\omega_{10} - \omega_{12})e^{i\omega_{10}t} = ib(\omega_{12} - \omega_{11})e^{i\omega_{11}t} \tag{34.7}$$

This equation is satisfied identically for all values of t. But this is only possible if

$$\omega_{10} = \omega_{11} \tag{34.8}$$

Similarly, substituting in (34.6) the value of $be^{i\omega_{11}t}$ from (34.5) and repeating the same argument, we obtain

$$\omega_{10} = \omega_{12} \tag{34.9}$$

Thus, the frequency of the wave does not change under reflection and refraction

$$\omega_{11} = \omega_{12} = \omega_{10} \tag{34.10}$$

We now show that the incident, reflected, and refracted rays lie in the same plane. In the boundary condition (34.4), $\mathbf{r}$ is the radius vector of a point on the boundary surface. We shall take one of the points of this surface as the origin. Then, $\mathbf{r}$ in (34.4) lies completely on the boundary surface. In this case, the boundary condition (34.4) may be written

$$a'e^{-i\mathbf{k}_{10}\cdot\mathbf{r}} + b'e^{-i\mathbf{k}_{11}\cdot\mathbf{r}} = c'e^{-i\mathbf{k}_{12}\cdot\mathbf{r}} \tag{34.11}$$

where a', b', c' are independent of $\mathbf{r}$. Applying the operation

$$\mathbf{r}\cdot\nabla = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}$$

to both sides of (34.11), and taking into account the fact that

$$\mathbf{r}\cdot\nabla e^{-i\mathbf{k}\cdot\mathbf{r}} = -i(\mathbf{k}\cdot\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}$$

we obtain

$$-ia'(\mathbf{k}_{10}\cdot\mathbf{r})e^{-i\mathbf{k}_{10}\cdot\mathbf{r}} - ib'(\mathbf{k}_{11}\cdot\mathbf{r})e^{-i\mathbf{k}_{11}\cdot\mathbf{r}} = -ic'(\mathbf{k}_{12}\cdot\mathbf{r})e^{-i\mathbf{k}_{12}\cdot\mathbf{r}} \tag{34.12}$$

Eliminating $c'e^{-i\mathbf{k}_{12}\cdot\mathbf{r}}$ from the right-hand side of (34.12), using (34.11), we obtain

$$ia'\{(\mathbf{k}_{10}\cdot\mathbf{r}) - (\mathbf{k}_{12}\cdot\mathbf{r})\}e^{-i\mathbf{k}_{10}\cdot\mathbf{r}} = ib'\{(\mathbf{k}_{12}\cdot\mathbf{r}) - (\mathbf{k}_{11}\cdot\mathbf{r})\}e^{-i\mathbf{k}_{11}\cdot\mathbf{r}} \tag{34.12a}$$

This equation is valid for all vectors $\mathbf{r}$ lying in the boundary plane, which is possible only if

$$\mathbf{k}_{10}\cdot\mathbf{r} = \mathbf{k}_{11}\cdot\mathbf{r} \tag{34.13}$$

Substituting from (34.11) for $b'e^{-i\mathbf{k}_{11}\cdot\mathbf{r}}$ in (34.12) and going through a similar argument, we obtain

$$\mathbf{k}_{10}\cdot\mathbf{r} = \mathbf{k}_{12}\cdot\mathbf{r} \tag{34.14}$$

Thus, we have

$$\mathbf{k}_{11}\cdot\mathbf{r} = \mathbf{k}_{12}\cdot\mathbf{r} = \mathbf{k}_{10}\cdot\mathbf{r} \tag{34.15}$$

Hence, it follows that $\mathbf{k}_{10}$, $\mathbf{k}_{11}$ and $\mathbf{k}_{12}$ are coplanar. In fact, $\mathbf{r}$ lies on the plane separating the media, and apart from that, it is arbitrary. We shall choose it in the direction perpendicular to one of the wave vectors, for example, perpendicular to $\mathbf{k}_{10}$. Then the condition (34.15) becomes

$$\mathbf{k}_{10}\cdot\mathbf{r} = 0 = \mathbf{k}_{11}\cdot\mathbf{r} = \mathbf{k}_{12}\cdot\mathbf{r}$$

But this means that $\mathbf{k}_{11}$ and $\mathbf{k}_{12}$ are also perpendicular to $\mathbf{r}$, i.e., they are coplanar with $\mathbf{k}_{10}$. Thus, it has been proved that the incident, reflected, and refracted rays are coplanar.

Relationship Between the Angles of Incidence, Reflection, and Refraction. Snell's Law. Let us take the origin of coordinates on the surface separating two dielectrics at the point of incidence of the ray. We shall take the xz plane as the plane of the incident, reflected and refracted rays. The z axis is perpendicular to the boundary surface and the x axis lies along this boundary (Fig. 41). Let $\mathbf{k}_{10}^{(0)}$, $\mathbf{k}_{11}^{(0)}$ and $\mathbf{k}_{12}^{(0)}$ be unit vectors in the directions of the corresponding rays. The various angles are shown in Fig. 41.

Equations (34.15) hold in any arbitrary coordinate system having the origin at a point of the boundary between the media. Let us now choose the origin at some point on the negative x axis (Fig. 41). In this case the direction of $\mathbf{r}$ is along the positive x axis. Hence, we have

$$(\mathbf{k}_{10}\cdot\mathbf{r}) = k_{10}r\cos\alpha_{10} \qquad (\mathbf{k}_{11}\cdot\mathbf{r}) = k_{11}r\cos\alpha_{11} \qquad (\mathbf{k}_{12}\cdot\mathbf{r}) = k_{12}r\cos\alpha_{12}$$

Therefore, equation (34.15) takes the form

$$k_{10}\cos\alpha_{10} = k_{11}\cos\alpha_{11} = k_{12}\cos\alpha_{12} \tag{34.16}$$

Let us denote the velocities of the incident, reflected, and refracted waves by v_{10}, v_{11}, and v_{12}, respectively. These velocities are related to the wave numbers k_{10}, k_{11}, and k_{12} by the equations

$$k_{10} = \frac{\omega}{v_{10}} \qquad k_{11} = \frac{\omega}{v_{11}} \qquad k_{12} = \frac{\omega}{v_{12}} \tag{34.17}$$

where the frequency invariance is taken into account. Since the incident and reflected waves are propagated in the same medium, we may write

$$v_{10} = v_{11} \qquad k_{10} = k_{11}$$

Consequently, equation (34.16) gives: $\cos\alpha_{10} = \cos\alpha_{11}$, $\alpha_{10} = \alpha_{11}$. Hence

$$\Theta_{10} = \Theta_{11} \tag{34.18}$$

This means that the angle of incidence is equal to the angle of reflection.

Furthermore, from (34.16) and (34.17), it follows that

$$\frac{1}{v_{10}}\cos\alpha_{10} = \frac{1}{v_{12}}\cos\alpha_{12} \tag{34.19}$$

Using the relationship $\cos \alpha_{10} = \sin \theta_{10}$ and $\cos \alpha_{12} = \sin \theta_{12}$, we may write (34.19) in the form

$$\frac{\sin \Theta_{10}}{\sin \Theta_{12}} = \frac{v_{10}}{v_{12}} \tag{34.20}$$

Since

$$v_{10} = \frac{1}{\sqrt{\epsilon_1 \mu_1}} \qquad v_{12} = \frac{1}{\sqrt{\epsilon_2 \mu_2}}$$

we may write (34.20) in the form

$$\frac{\sin \Theta_{10}}{\sin \Theta_{12}} = \sqrt{\frac{\epsilon_2 \mu_2}{\epsilon_1 \mu_1}} = n_{12} \tag{34.21}$$

i.e., the ratio of the sine of the angle of incidence to the sine of the angle of refraction equals the ratio of the refractive index of the second medium with respect to the first (*Snell's law*).

Relationship Between the Intensities of the Incident, Reflected, and Refracted Waves. Fresnel's Formulas. Let us consider the case of normal incidence of waves on the boundary separating two media. We shall take the z axis along the direction of propagation of the incident wave, perpendicular to the boundary surface. Let the x and y axes lie in the boundary plane (Fig. 42) and the x axis coincide with the direction of **E** for the incident

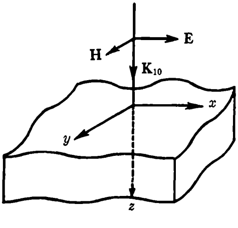


Fig. 42

wave. It follows that the y axis coincides with the direction of **H**. Thus, for the incident wave, we may write down the equations

$$E_x = E_{10}^{(0)} e^{i(\omega t - k_{10} z)} \qquad E_y = E_z = 0 \tag{34.22}$$

$$H_y = \sqrt{\frac{\epsilon_1}{\mu_1}} E_{10}^{(0)} e^{i(\omega t - k_{10} z)} \qquad H_x = H_z = 0 \tag{34.23}$$

In (34.23) we use the relationship (32.22) between the amplitudes of **H**

and **E** in a plane wave. For the refracted wave, the expressions for the vectors may be written

$$E_x = E_{12}^{(0)} e^{i(\omega t - k_{12} z)} \qquad E_y = E_z = 0 \tag{34.24}$$

$$H_y = \sqrt{\frac{\epsilon_2}{\mu_2}}\, E_{12}^{(0)} e^{i(\omega t - k_{12} z)} \qquad H_x = H_z = 0 \tag{34.25}$$

To formulate the expressions for **E** and **H** in the reflected wave, we must take two facts into account. First, the reflected wave travels in the negative z direction. Second, since the vectors **E**, **H**, and the direction vector of the wave form a right-hand screw triad of mutually perpendicular vectors, the direction of propagation of the wave may be reversed only if either **E** or **H** is reversed, while the direction of the other vector remains the same. In the reflected wave, the direction of **H** is reversed. Taking these two facts into account, we find that the field vectors for the reflected wave take the form

$$E_x = E_{11}^{(0)} e^{i(\omega t + k_{11} z)} \qquad E_y = E_z = 0 \tag{34.26}$$

$$H_y = -\sqrt{\frac{\epsilon_1}{\mu_1}}\, E_{11}^{(0)} e^{i(\omega t + k_{11} z)} \qquad H_x = H_z = 0 \tag{34.27}$$

The boundary conditions expressing the continuity of the tangential components of **E** and **H** are

$$E_{10}^{(0)} + E_{11}^{(0)} = E_{12}^{(0)} \tag{34.28}$$

$$\sqrt{\epsilon_1}\, E_{10}^{(0)} - \sqrt{\epsilon_1}\, E_{11}^{(0)} = \sqrt{\epsilon_2}\, E_{12}^{(0)} \tag{34.29}$$

where we have assumed, for dielectrics

$$\mu_1 \approx \mu_2 \approx \mu_0$$

The solutions of the system of algebraic equations (34.28) and (34.29) are written in the form

$$E_{12}^{(0)} = \frac{2}{1 + n_{12}} E_{10}^{(0)} \tag{34.30}$$

$$E_{11}^{(0)} = \frac{1 - n_{12}}{1 + n_{12}} E_{10}^{(0)} \tag{34.31}$$

where $n_{12} = \sqrt{\epsilon_2/\epsilon_1}$ is the refractive index of the second medium with respect to the first.

The intensity of the wave is described by the absolute magnitude of the Poynting vector, which in this case is written

$$|\mathbf{P}| = \sqrt{\frac{\epsilon}{\mu}}\, E^2$$

where $\mu \approx \mu_0$. Since **E** and **H** vary harmonically, the mean values per period of these vectors are related to the amplitudes E_0 and H_0 by the equations

$$\langle E^2 \rangle = E_0^2 \langle \cos \omega t \rangle = \frac{1}{2} E_0^2 \qquad \langle H^2 \rangle = \frac{1}{2} H_0^2$$

Consequently, S, the mean intensity of the wave per period, is related to the amplitude by the equation

$$\langle |\mathbf{P}| \rangle = S = \frac{1}{2} \sqrt{\frac{\epsilon}{\mu}} E_0^2$$

Thus, the mean intensities per period of the incident, reflected, and refracted waves are

$$S_{10} = \frac{1}{2} \sqrt{\frac{\epsilon_1}{\mu_0}} E_{10}{}^{(0)2} \tag{34.32}$$

$$S_{11} = \frac{1}{2} \sqrt{\frac{\epsilon_1}{\mu_0}} E_{11}{}^{(0)2} = \frac{1}{2} \sqrt{\frac{\epsilon_1}{\mu_0}} \left(\frac{1 - n_{12}}{1 + n_{12}} \right)^2 E_{10}{}^{(0)2} = \left(\frac{1 - n_{12}}{1 + n_{12}} \right)^2 S_{10} \tag{34.33}$$

$$S_{12} = \frac{1}{2} \sqrt{\frac{\epsilon_2}{\mu_0}} E_{12}{}^{(0)2} = \frac{1}{2} \sqrt{\frac{\epsilon_2}{\mu_0}} \left(\frac{2}{1 + n_{12}} \right)^2 E_{10}{}^{(0)2} = \frac{4n_{12}}{(1 + n_{12})^2} S_{10} \tag{34.34}$$

Equations (34.32) to (34.34) are called *Fresnel's formulas*. They describe the relationship between the intensities of the incident, reflected, and refracted waves. The ratio of the intensity of the reflected wave to that of the incident wave r is called the *coefficient of reflection*. From equation (34.33) we obtain

$$r = \frac{S_{11}}{S_{10}} = \left(\frac{1 - n_{12}}{1 + n_{12}} \right)^2 \tag{34.35}$$

The *coefficient of refraction* χ_{ref} is defined in a similar manner as the ratio of the intensity of the refracted wave to that of the incident wave

$$\chi_{\text{ref}} = \frac{S_{12}}{S_{10}} = \frac{4n_{12}}{(1 + n_{12})^2} \tag{34.36}$$

Equations (34.32) to (34.34) show that

$$S_{11} + S_{12} = S_{10} \tag{34.37}$$

This means that the energy of the incident wave is completely transformed into the energy of the reflected and refracted waves, i.e., in the case of reflection and refraction at a boundary separating dielectrics, there is no transformation of the energy of the electromagnetic field into other forms of energy.

§35. Motion of Electromagnetic Waves along Transmission Lines

From the law of conservation of energy, it follows that the work done in some region of space by the electromagnetic field may be performed either at the expense of a reduction of the energy of the field, or by the flow of electromagnetic energy into that region. The electromagnetic energy flux is given by the Poynting vector

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \tag{35.1}$$

If the energy is transmitted along conductors, then this energy can only be transmitted by the motion of the electromagnetic field associated with the electric current, with the conductor acting as a guide along which the energy moves.

Let us consider a long straight conductor, of radius r, carrying a steady current of density j (Fig. 43).

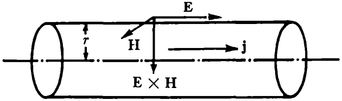


Fig. 43

In the absence of applied forces $\mathbf{j} = \lambda\mathbf{E}$, and, therefore, the electric field intensity is directed along the axis of the conductor and is equal to

$$\mathbf{E} = \frac{\mathbf{j}}{\lambda} \tag{35.2}$$

From the boundary condition expressing the continuity of the tangential components of the electric field, the same field also exists outside the conductor.

The magnetic field close to the surface of the conductor is perpendicular to $\mathbf{j}$, and is directed along the tangent to the conductor surface. Its absolute magnitude is

$$H = \frac{j\pi r^2}{2\pi r} = \frac{j}{2} r \tag{35.3}$$

Thus, the Poynting vector (35.1) is directed along the radius toward the center of the conductor and equals

$$P = EH = \frac{1}{2\lambda} j^2 r \tag{35.4}$$

Hence, the electromagnetic energy flows into the conductor from the surrounding space through the lateral surface. For a section of the conductor

of length l, the amount of energy flowing into the conductor per unit time is

$$Q = P\,2\pi rl = \frac{j^2}{\lambda}\pi r^2 l \tag{35.5}$$

On the other hand, it is known that according to the Joule-Lenz law, the amount of heat liberated by a length l of a conductor is

$$Q_j = \frac{j^2}{\lambda}\pi r^2 l \tag{35.6}$$

Comparison of (35.5) with (35.6) shows that the whole energy liberated in the form of heat on passage of an electric current enters the conductor from the surrounding space. This example indicates that the energy transmitted by a current moves principally in the space surrounding the conductor. The conductor plays the part of a guide along which the energy is transmitted. This is especially clear in the transmission of energy by means of a cable (Fig. 44), which we shall discuss now.

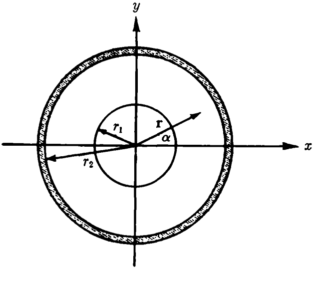


Fig. 44

The current moves along the inner conductor in one direction, and returns along the outer conductor in the opposite direction. The space between the inner and outer conductors is filled with a dielectric. In order to simplify the problem, let us suppose that the cable has no electrical resistance so that there is no loss of energy during transmission along it. The entire drop in potential occurs at the point where energy is consumed. Let this potential drop be V. Then the difference in potential between the outer and inner conductors is V. It follows that an electric field acts

between the inner and outer conductors. Due to the axial symmetry of the problem, the component E_α of this field is equal to zero. The lines of force of the magnetic field are concentric circles with their centers on the axis of the cable. Since there are no losses in the cable, there is no flux of energy into the inner and outer conductors from the space between them, i.e., the radial component of the Poynting vector equals zero. This means that $E_z = 0$. Maxwell's equation

$$\operatorname{div} \mathbf{E} = \frac{1}{r}\frac{\partial}{\partial r}(rE_r) = 0 \tag{35.7}$$

then gives the following expression for E_r

$$E_r = \frac{a_0}{r} \tag{35.8}$$

The value of a_0 is defined by the difference in potential between the inner and outer conductors

$$V = \int_{r_1}^{r_2} E_r\,dr = a_0 \int_{r_1}^{r_2} \frac{dr}{r} = a_0 \ln\frac{r_2}{r_1}$$

Consequently

$$E_r = \frac{V}{\ln\dfrac{r_2}{r_1}}\frac{1}{r} \tag{35.9}$$

On the other hand, the magnetic field in the cable is

$$H_\alpha = \frac{I}{2\pi r} \tag{35.10}$$

and, hence, the Poynting vector is given by

$$P_z = E_r H_\alpha = \frac{1}{2\pi}\frac{VI}{\ln\dfrac{r_2}{r_1}}\frac{1}{r^2} \tag{35.11}$$

Thus, the energy is transmitted along the cable in the space between the inner and outer conductors. The density of the energy flux is given by (35.11). Outside the cable, as shown in Problem 3 of Chapter 3, the field is equal to zero and there is no electromagnetic energy. P_z, therefore, represents the total flow of the electromagnetic energy. The electromagnetic energy passing through a cross section of the cable per unit time is

$$Q = \int_S P_z\,dS = \frac{1}{2\pi}\frac{VI}{\ln\dfrac{r_2}{r_1}}\int_0^{2\pi} d\alpha \int_{r_1}^{r_2}\frac{dr}{r} = VI \tag{35.12}$$

Also, it is well known that for a current I and a potential difference V the total power used by a consumer is

$$Q_p = IV \tag{35.13}$$

Comparison of (35.12) with (35.13) shows that the total energy used by the consumer is transmitted along the cable in the space between the inner and outer conductors.

In the case of a low-frequency alternating current, there are no differences in principle from the description given above. If the direction of the current in the cable is reversed, then the directions of E_r and H_α are reversed, but the direction of the Poynting vector remains the same. Hence, although the direction of the current is reversed, the direction of motion of the electromagnetic energy remains the same, moving, as before, to the consumer.

In other transmission lines, the transmission takes place, in principle, according to this simplified description, but the configuration of the electromagnetic fields is more complicated.

At high frequencies, when the wavelength is of the order of the distance between the conductors, the transmission line begins to radiate electromagnetic energy and acts as an aerial. Furthermore, at very high frequencies large losses of energy occur in the insulation of conductors. Ordinary transmission lines, are, therefore, unsuitable for transmitting energy at high frequencies corresponding to wavelengths in the centimeter range. In this case it is necessary to transmit the electromagnetic energy along waveguides.

In the simplest case, a waveguide consists of a hollow straight metal tube, of constant cross section along the length. The theory of propagation of waves along waveguides is developed from Maxwell's theory. We shall not describe this theory, but shall only state two important properties of waveguides.

Let the z axis lie along the axis of a waveguide, which has an arbitrary cross section. The walls of the waveguide are assumed to be perfectly conducting. In such a waveguide, all possible electromagnetic waves may be divided into two types: those with $H_z = 0$, and those with $E_z = 0$. The waves in which $H_z = 0$ are called E-waves or, *electric modes*, and the waves in which $E_z = 0$ are called H-waves, or *magnetic modes*.

The wavelength of waves which can be propagated along a waveguide must not exceed a certain maximum value. If the wavelength exceeds this value, the waves cannot travel along the waveguide. This maximum permissible wavelength is of the order of the cross section of the waveguide. Hence, waveguides are especially important in ultrashort-wave radio ap-

paratus for transmitting the electromagnetic energy from a generator of oscillations to a transmitter.

PROBLEMS

1 Air becomes ionized when the electric field intensity reaches $E \approx 30$ kV/cm. Determine the density of the flux of energy of low frequency plane electromagnetic waves sufficient to ionize the air.
Hint: The amplitude of the electric field vector must equal the ionization intensity.

$$\textit{Answer:}\quad \langle P\rangle = \frac{1}{2}\sqrt{\frac{\epsilon_0}{\mu_0}}\,E_0{}^2 \approx 1.2 \times 10^{10}\ \mathrm{W/m^2} = 1.2 \times 10^3\ \mathrm{k\ W/cm^2}$$

2 A plane-polarized electromagnetic wave of angular frequency $\omega = 10^6\ \mathrm{sec}^{-1}$ is transmitted from a straight segment to a conductor loop. The vector **H** of the wave is perpendicular to the plane of the loop. The linear dimensions of the loop are small compared with the wavelength. The area of the loop is $S = 100$ cm^2, the mean energy flux density in the wave is $\langle P\rangle = 1$ W/m^2. Find the maximum induced emf in the circuit.
Hint: Use the law of electromagnetic induction.

$$\textit{Answer:}\quad \mathcal{E}^{\mathrm{ind}}_{\mathrm{max}} = \sqrt{2\langle P\rangle\mu_0}\,(\epsilon_0\mu_0)^{1/4}S\omega \approx 9 \times 10^{-3}\ \mathrm{V}$$

PART II

Electron Theory

IT IS well known that the electron carries a single elementary negative charge, and the proton carries a single elementary positive charge. The charges on proton and electron are equal in magnitude and opposite in sign. The absolute value of such a charge is $|e| = 1.6 \times 10^{-19}$ coul.

In modern physics, the electron is considered as a point particle. The proton may also be considered as a point particle in the electron theory, although in recent years the electromagnetic structure of the proton has been investigated, and it has been shown that the charge of the proton is distributed in a region with a root mean square radius of the order of 10^{-13} cm.

No charges exist in nature which are fractions of the elementary charge e. Every charge must be equal to an integral multiple of the elementary charge. Nor can a charge be distributed continuously in space; it consists of an assembly of elementary point charges. In the first part of this book, we ignored this discrete structure of the electric charge. Since the elementary charge is small, it is permissible to do this in many cases without committing a serious error. But this is not always possible. In the second part of this book, we shall take into account the existence of elementary point carriers of charge: the *electron* and the *proton*.

Another point which we have ignored in the first part is the atomic structure of matter. We have considered matter to be continuous, and the electrical properties of matter to be described by macroscopic physical parameters. Such an approach is permissible in many cases, but not in all. In the second part of this book, the atomic structure of matter will be taken into account. Hence, many phenomena will be explained which are inexplicable from the point of view of phenomenological electrodynamics, e.g., the dispersion of light.

CHAPTER 7

Interaction of Charges with the Electromagnetic Field

§36. Fundamental Equations of Electron Theory

Phenomenological electrodynamics allows for the properties of matter by introducing the coefficients ϵ, μ, and λ, which describe the properties of a medium. To describe electromagnetic phenomena in matter, the electron theory does not use macroparameters, except the permittivity of empty space ϵ_0 and the permeability of empty space μ_0. Hence, the electron theory is based on the field vectors in empty space, which we shall denote by small letters, **e**, **d**, **h**, and **b**. According to the electron theory, all electromagnetic phenomena in material media and all electric and magnetic properties of matter may be explained by laws of the creation of an electromagnetic field by charges and currents in empty space. Hence, the effect of a material medium on the electromagnetic field is reduced to the effect of additional charges and currents which appear in the medium under actual conditions.

It is now known that not all electromagnetic properties of matter may be explained in this way. For example, we cannot explain in this way the magnetic moment of elementary particles, which plays an important part in the explanation of the magnetic properties of matter, nor can we explain the quantum laws, etc. Nevertheless, within the framework of the electron theory, we can explain a wide range of phenomena which are inexplicable from the point of view of phenomenological electrodynamics.

As stated, the electromagnetic field vectors in the electron theory will

be denoted by small letters. They obey, like the electromagnetic field vectors *in vacuo*, the relationship

$$\mathbf{d} = \epsilon_0 \mathbf{e}$$
$$\mathbf{b} = \mu_0 \mathbf{h}$$

We use a special notation for these vectors for a good reason. They differ from the electromagnetic field vectors *in vacuo*, which are the basis of phenomenological electrodynamics. In phenomenological electrodynamics, we ignore the atomic structure of matter, although it is known that the electrons in atoms are in rapid motion, and that the charge density varies greatly from point to point, and, hence, **e**, **d**, **h**, and **b**, generally speaking, vary rapidly in time, as well as from point to point. **E**, **D**, **H**, and **B** in phenomenological electrodynamics do not allow for these rapid changes which are due to the atomic structure of matter and the discrete nature of the charges; **E**, **D**, **H**, and **B** are the average "smoothed-out" values.

A current is produced by the motion of charge. If the charge density is ρ and the velocity of the charges is $\mathbf{v}$, then, by definition of the current density $\mathbf{j}$, we have

$$\mathbf{j} = \rho \mathbf{v} \tag{36.1}$$

This equation holds both for a continuous distribution of charge and for discrete point charges. In the case of a point charge e at a point $\mathbf{r}_0$, equation (1.8) for ρ becomes

$$\rho = e\delta(\mathbf{r} - \mathbf{r}_0) \tag{36.2}$$

where $\delta(\mathbf{r} - \mathbf{r}_0)$ is the δ-function defined by equations (15.17a) and (15.17b).

The fundamental equations of the electron theory are called the *Maxwell-Lorentz equations*

$$\left.\begin{aligned} \operatorname{curl} \mathbf{h} &= \rho\mathbf{v} + \frac{\partial \mathbf{d}}{\partial t} && \text{(I)} \\ \operatorname{curl} \mathbf{e} &= -\frac{\partial \mathbf{b}}{\partial t} && \text{(II)} \\ \operatorname{div} \mathbf{b} &= 0 && \text{(III)} \\ \operatorname{div} \mathbf{d} &= \rho && \text{(IV)} \end{aligned}\right\} \tag{36.3}$$

where

$$\mathbf{d} = \epsilon_0 \mathbf{e} \qquad \mathbf{b} = \mu_0 \mathbf{h} \qquad \text{(V)} \tag{36.3a}$$

The equation of continuity, which expresses the law of conservation of charge, takes the form

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \rho\mathbf{v} = 0 \tag{36.4}$$

The volume density of the force acting on the charges is given by

$$\mathbf{f} = \rho(\mathbf{e} + \mathbf{v} \times \mathbf{b}) \tag{36.5}$$

The equation for the energy density of the magnetic field and the density of the energy flux are also completely analogous to phenomenological electrodynamics

$$u = \frac{1}{2}(\mathbf{e}\cdot\mathbf{d} + \mathbf{h}\cdot\mathbf{b}) \tag{36.6}$$

$$\mathbf{P} = \mathbf{e} \times \mathbf{h} \tag{36.7}$$

The electromagnetic field also has momentum, the density of the momentum $\mathfrak{g}$ being given by

$$\mathfrak{g} = \frac{1}{c^2}\mathbf{P} = \frac{1}{c^2}\mathbf{e} \times \mathbf{h} \tag{36.8}$$

This equation will be proved in §41. If an electromagnetic wave is absorbed by a body, or reflected from it, then a corresponding momentum is transmitted to the body and appears in the form of pressure of the electromagnetic wave.

The physical meaning of the Maxwell-Lorentz equations is the same as that of Maxwell's equations for empty space, but their theoretical significance is different. Maxwell's equations for empty space are a special case of Maxwell's equations for a medium with properties described by ϵ, μ, λ. But the Maxwell-Lorentz equations are the fundamental equations of the electron theory, and are suitable for describing phenomena both in empty space and in a medium. The effect of a medium on the field is determined by the actual nature of the distribution and motion of the charges in the medium, including those in atoms and molecules. In the equations of the electron theory, this is allowed for by the quantities $\mathbf{v}$ and ρ. If the electromagnetic field vectors in Maxwell's equations for empty space are taken to mean the true vectors of the electromagnetic field, formulated with the motion of charges within atoms and molecules taken into account, then these field vectors will be identical with the vectors of the electron theory denoted in the foregoing discussion by small letters. In order not to complicate the notation, we shall, in future, use capital letters to denote the field vectors, except where this would lead to misunderstandings.

§37. Motion of an Electron in an Electromagnetic Field

Equation of Motion. When a point charge e moves in an electromagnetic field at a velocity $\mathbf{v}$, the Lorentz force

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{37.1}$$

acts upon it. Hence, Newton's equation of motion takes the form

$$\frac{d}{dt}(m\mathbf{v}) = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{37.2}$$

Throughout Part II of the present book, we shall assume that the velocity of a particle is much less than the velocity of light. Then the mass may be taken to be constant. Motion at velocities close to the velocity of light will be discussed in Part III.

Motion in a Magnetic Field. In this case, the equation of motion takes the form

$$m\frac{d\mathbf{v}}{dt} = e\mathbf{v} \times \mathbf{B} \tag{37.3}$$

Taking the scalar product of both sides by $\mathbf{v}$, we obtain

$$m\frac{d\mathbf{v}}{dt} \cdot \mathbf{v} = e(\mathbf{v} \times \mathbf{B} \cdot \mathbf{v}) = 0$$

since the scalar triple product with two of the vectors parallel is equal to zero. Thus

$$m\frac{d\mathbf{v}}{dt} \cdot \mathbf{v} = \frac{d}{dt}\left(\frac{m\mathbf{v}^2}{2}\right) = 0$$

and, therefore

$$\frac{mv^2}{2} = \text{const} \qquad v^2 = \text{const} \tag{37.4}$$

Equations (37.4) state that the magnetic field performs no work and produces no change in the kinetic energy of the particle nor in its speed, because the force due to the magnetic field is always perpendicular to the direction of motion. The force is, therefore, perpendicular to the displacement, and, hence, does no work. The force due to the magnetic field only changes the direction of motion of the particle.

Let the magnetic field be homogeneous and constant with respect to time. Then the velocity of the particle $\mathbf{v}$ may be represented in the form of a sum of two velocities

$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel \tag{37.5}$$

where $\mathbf{v}_\perp$ is the component of velocity perpendicular to the magnetic field, and $\mathbf{v}_\parallel$ is the component of velocity in the direction of the magnetic field.

The component of the Lorentz force in the direction of the magnetic field is equal to zero

$$\mathbf{F}_{\parallel} = e(\mathbf{v} \times \mathbf{B})_{\parallel} = 0 \tag{37.6}$$

Taking the projections of both sides of (37.3) along the direction of the magnetic field, and using (37.6), we obtain

$$m\frac{d\mathbf{v}_{\parallel}}{dt} = 0 \qquad \mathbf{v}_{\parallel} = \text{const} \tag{37.7}$$

i.e., the velocity of the particle is constant in the direction of the magnetic field.

Substituting (37.5) in (37.3), using (37.7) and the equation

$$\mathbf{v} \times \mathbf{B} = (\mathbf{v}_{\perp} + \mathbf{v}_{\parallel}) \times \mathbf{B} = \mathbf{v}_{\perp} \times \mathbf{B}$$

we obtain the following equation for the transverse component of velocity $\mathbf{v}_{\perp}$

$$m\frac{d\mathbf{v}_{\perp}}{dt} = e\mathbf{v}_{\perp} \times \mathbf{B} \tag{37.8}$$

To solve this equation, we recall that

$$v_{\perp}^2 + v_{\parallel}^2 = v^2 = \text{const}$$

Hence, from (37.7), we have

$$v_{\perp}^2 = \text{const}$$

The angle between $\mathbf{v}_{\perp}$ and $\mathbf{B}$ in equation (37.8) remains constant throughout the motion and equal to $\pi/2$; $\mathbf{v}_{\perp}$ and $\mathbf{B}$ do not change in magnitude. Hence, the force in (37.8) is constant in magnitude and is perpendicular to the velocity. Consequently, equation (37.8) describes the motion with constant acceleration that is always perpendicular to the velocity. This is motion on a circle. Denoting the radius of the circle by r, we may write the equation of motion using projections along the normal to the trajectory in the following form

$$\frac{mv_{\perp}^2}{r} = |e|v_{\perp}B \tag{37.9}$$

The angular frequency of revolution ω is equal to

$$\omega = \frac{2\pi}{T} = \frac{2\pi}{\left(\frac{2\pi r}{v_{\perp}}\right)} = \frac{v_{\perp}}{r}$$

Hence, equation (37.9) gives the following important relationships

$$\omega = \frac{|e|B}{m} \qquad r = \frac{v_{\perp}}{\omega} = \frac{mv_{\perp}}{|e|B} \tag{37.10}$$

Motion in an Electric Field That Is Constant in Time. Let us consider the

motion of a charge e in an electric field $\mathbf{E} = -\operatorname{grad} \varphi$ that is constant in time, in the absence of a magnetic field.

In this case, the equations of motion take the following form

$$m \frac{d\mathbf{v}}{dt} = -e \operatorname{grad} \varphi \tag{37.11}$$

Taking the scalar product of both sides of this equation with $\mathbf{v} = d\mathbf{r}/dt$, we obtain

$$m \frac{d\mathbf{v}}{dt} \cdot \mathbf{v} = \frac{d}{dt}\left(\frac{mv^2}{2}\right) = -e \operatorname{grad} \varphi \frac{d\mathbf{r}}{dt} = -e \frac{d\varphi}{dt}$$

Hence, we obtain the equation

$$\frac{mv^2}{2} + e\varphi = \text{const} \tag{37.12}$$

which expresses the law of conservation of energy for motion in an electrostatic field.

Thus, if an electron is initially in a state of rest and is then subjected to a potential difference V, we may write, on the basis of (37.12)

$$\frac{mv^2}{2} = |e| V$$

Hence, it follows that

$$v = \sqrt{\frac{2|e|}{m} V} \tag{37.13}$$

For an electron, equation (37.13) takes the following form

$$v \approx 6 \times 10^5 \sqrt{V} \text{ m/sec} = 600\sqrt{V} \text{ km/sec} \tag{37.14}$$

Thus, the application of a potential difference of only 1 V to an electron gives the latter a velocity of 600 km/sec.

In atomic physics, energy is measured in electron volts. One electron volt is the energy acquired by a particle carrying a charge equal in magnitude to the electron charge, on being subjected to a potential difference of one volt

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ coul V} = 1.6 \times 10^{-19} J \tag{37.15}$$

Drift of Charged Particles in Crossed Fields. Let us consider the case when the radius of curvature of the trajectory of a particle is much smaller than the linear dimensions of the region of motion, so that in the course of its motion, the particle makes many revolutions. Let the motion take place in homogeneous crossed fields $\mathbf{E} \perp \mathbf{B}$ (Fig. 45).

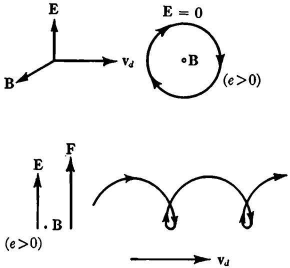


Fig. 45

We shall seek the solution of the equation of motion (37.2) in the form

$$\mathbf{v} = \mathbf{v}' + \frac{1}{B^2}\mathbf{E} \times \mathbf{B} \tag{37.16}$$

Substituting this expression for $\mathbf{v}$ in (37.2), we obtain

$$m\frac{d\mathbf{v}'}{dt} = e\mathbf{E} + e\mathbf{v}' \times \mathbf{B} + \frac{1}{B^2}\{\mathbf{B}\cdot\mathbf{E}\mathbf{B} - \mathbf{E}B^2\} = e\mathbf{v}' \times \mathbf{B}$$

where we have used the formula for the expansion of the vector triple product, and recalling that the scalar product of $\mathbf{E}$ and $\mathbf{B}$ is zero, since these vectors are orthogonal. The equation for $\mathbf{v}'$

$$m\frac{d\mathbf{v}'}{dt} = e\mathbf{v}' \times \mathbf{B}$$

is identical with (37.8), i.e., $\mathbf{v}'$ describes rotational motion. Thus, the motion in crossed fields given by (37.16), may be described as follows. The particle describes a trajectory with velocity $\mathbf{v}'$ about a center moving with a velocity $\mathbf{v}_d$

$$\mathbf{v}_d = \frac{1}{B^2}\mathbf{E} \times \mathbf{B} \tag{37.17}$$

As a result of these two motions, the particle is displaced in the direction of the velocity $\mathbf{v}_d$. This motion is called the *drift* of the particle, and $\mathbf{v}_d$ is called the *drift velocity*. In absolute magnitude, the drift velocity equals

$$v_d = \frac{1}{B^2}|\mathbf{E} \times \mathbf{B}| = \frac{E}{B} \tag{37.18}$$

This equation may be used only when $v_d \ll c$. It must also be observed

that the direction of drift is independent of the sign of the charge on the particle.

The physical picture of drift is as follows (Fig. 45). In the absence of an electric field, a charged particle in a magnetic field moves in a circle at a constant velocity. When an electric field is applied perpendicular to the magnetic field, the velocity of the particle becomes variable. When the particle is moving in the direction of the electric field, the velocity of the particle increases and, hence, the radius of curvature of the trajectory also increases (the upper part of the trajectory in Fig. 45). Having changed the direction of motion to one with a large radius of curvature, the particle begins to move against the electric field, and, hence, the velocity and radius of curvature of the particle decrease (lower part of the trajectory in Fig. 45). Consequently, a change in direction occurs over a smaller part of the trajectory. As a result, the particle drifts, as shown in Fig. 45. The velocity of drift is given by equations (37.17) and (37.18).

The fact that the direction of drift in crossed electric and magnetic fields is independent of the sign of the charge is due to the fact that when the sign of the charge changes, there is a simultaneous change in the direction of revolution and the direction of action of the electric field force. The physical picture of drift shows that if, instead of an electric force, some other force acts in the same direction, then the drift remains unchanged. If the direction of action of this force is independent of the sign of the charge, then particles of different charges will drift in different directions. But, if the direction of action of this force is reversed when the sign of the charge changes, then the direction of drift for unlike particles will be the same.

Drift of Particles in an Inhomogeneous Magnetic Field. Let us consider a magnetic field such that its magnitude varies in the direction perpendicular to the direction of the magnetic field (Fig. 46). If the field were homogeneous, then the particle would move in a circle. However, since it is inhomogeneous, the radius of curvature of the particle changes in the course of the motion; where the field is stronger, the radius of curvature is smaller, and vice versa. Thus, the picture for the case of crossed fields is completely repeated in this case. Consequently, the drift takes place in a direction that is perpendicular both to the magnetic field and to the direction along which the field is inhomogeneous. It follows directly from Fig. 46 that particles of different sign drift in different directions.

Adiabatic Invariance of the Magnetic Moment. In many cases of practical importance, a magnetic field changes slowly at distances of the order of the radius of curvature of the trajectory of a particle in a time of the order of one revolution period. In this case, by analogy with the magnetic

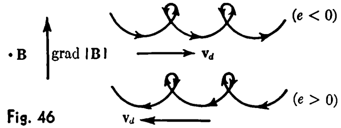


Fig. 46

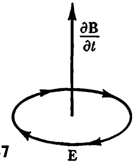


Fig. 47

moment of a circular current, we may introduce the concept of the magnetic moment of the particle, and show that this moment is conserved in slowly changing magnetic fields.

The magnetic moment M of a circular current I is, by definition

$$M = IS \tag{37.19}$$

where S is the area enclosed by the current. If a charge $|e|$ moves round a circle of radius r with a period of revolution T, then the magnetic moment associated with the motion of this particle, is, as defined by (37.19), equal to

$$M = \frac{|e|}{T}\pi r^2$$

Taking into account the fact that

$$T = \frac{2\pi r}{v_\perp} \qquad r = \frac{mv_\perp}{|e|B} \tag{37.20}$$

we obtain the following expression for the magnetic moment of the particle

$$M = \frac{\frac{1}{2}mv_\perp{}^2}{B} = \frac{W_\perp}{B} \tag{37.21}$$

where

$$W_\perp = \frac{m}{2}v_\perp{}^2$$

is the kinetic energy associated with the component of velocity of the particle perpendicular to the magnetic field.

We shall now show that the magnetic moment M of a particle is conserved in slowly changing magnetic fields. First, we shall consider its behavior in a field that varies with time (Fig. 47). By Faraday's law of electromagnetic induction, a particle moving in a circle experiences an induced electric field

$$E = \frac{1}{2\pi r}\frac{d\Phi}{dt} = \frac{r}{2}\frac{dB}{dt}$$

This electric field, in one revolution, transfers to the particle an energy

$$\Delta\left(\frac{1}{2}\,mv_\perp{}^2\right) = 2\pi rE|e| = |e|\pi r^2\,\frac{dB}{dt} \tag{37.22}$$

Since the period T is small in comparison with the time constant of the magnetic field, and the energy of the particle changes only slightly during one revolution, we may divide both sides of the equation by T and write

$$\frac{\Delta\left(\frac{1}{2}\,mv_\perp{}^2\right)}{T} \approx \frac{d}{dt}\left(\frac{1}{2}\,mv_\perp{}^2\right) = \frac{|e|\pi r^2}{T}\,\frac{dB}{dt} = M\,\frac{dB}{dt} \tag{37.23}$$

where (37.20) is taken into account. Using the definition of the magnetic moment, (37.21), we may write (37.23) in the form

$$\frac{d}{dt}\,(MB) = M\,\frac{dB}{dt} \tag{37.24}$$

Hence, it follows that

$$M = \text{const}$$

This proves that in a magnetic field which changes slowly with time, the magnetic moment of a moving charge conserves its magnitude. To prove the conservation of the magnetic moment in a field slowly changing in space, we shall consider the motion of a particle in the direction of the changing magnetic field (Fig. 48). If the magnetic field is increasing in the

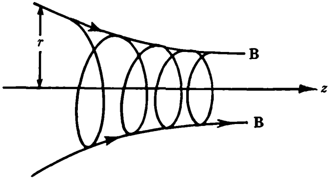


Fig. 48

direction of the z axis, then the lines of force converge in this direction. It follows directly from Fig. 48 that in this case, the magnetic lines of force have a component B_r. Due to the presence of the velocity $v_\perp$ and the component of the field B_r, a force

$$\mathbf{F} = e\mathbf{v}_\perp \times \mathbf{B}_r$$

acts on the charge, reducing the z component of the velocity of the particle,

i.e., the velocity suffers deceleration in the direction of the increasing magnetic field. To determine the value of this retarding force, we can use equation (22.22) for the force on a magnetic moment, taking (37.21) as the magnetic moment. Then, for the z component of the velocity, we may write down the equation

$$m\frac{dv_z}{dt} = -M\frac{\partial B_z}{\partial z} = -M\frac{\partial B}{\partial z} \tag{37.25}$$

since $B_z \approx B$. Multiplying both sides of this equation by v_z, and recalling that

$$\frac{\partial B}{\partial z}v_z = \frac{\partial B}{\partial z}\frac{dz}{dt} = \frac{dB}{dt}$$

we obtain

$$\frac{d}{dt}\left(\frac{mv_z^2}{2}\right) = -M\frac{dB}{dt} \tag{37.26}$$

But

$$\frac{mv_z^2}{2} + \frac{mv_\perp^2}{2} = \frac{mv^2}{2} = \text{const}$$

and, therefore, (37.26) becomes

$$\frac{d}{dt}\left(\frac{mv_\perp^2}{2}\right) = \frac{d}{dt}(MB) = M\frac{dB}{dt} \tag{37.27}$$

Hence, it follows that

$$M = \text{const} \tag{37.28}$$

i.e., the magnetic moment is conserved. This proves the adiabatic invariance of the magnetic moment, i.e., its conservation in a magnetic field slowly changing in both time and space.

The conservation of the magnetic moment means that the particle moves on the surface of a single magnetic tube of force, i.e., a tube having its surface formed by lines of force of the magnetic field. The magnetic flux passing through this tube is equal to

$$\Phi = \pi r^2 B = \frac{2\pi m}{e^2}\frac{W_\perp}{B} = \frac{2\pi m}{e^2}M = \text{const}\, M$$

Thus, the constancy of the magnetic moment is equivalent to the constancy of the magnetic flux enclosed by the orbit of the particle. But this means that the particle moves on the surface of a magnetic tube. In an inhomogeneous field, the particle moves from one tube of force to another, but in such a way that the magnetic fluxes enclosed by these tubes are identical.

Magnetic Mirrors. The fact that the magnetic moment of a moving charge is constant means that when it moves into a region of increasing magnetic induction, the value of $v_\perp^2$ increases. On the other hand, the square of the total velocity in a magnetic field is constant

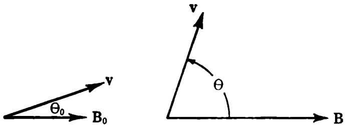
$$v_\perp^2 + v_\parallel^2 = v^2 = \text{const}$$

Consequently, the velocity in the direction of the magnetic field must decrease. If the increase in the magnetic field is sufficiently great, the velocity $v_\parallel$ will at some point become zero, and the particle will then move away in a different direction. Thus, a region where the magnetic field is strong reflects charged particles, i.e., it acts as a *magnetic mirror.*

Let the velocity $\mathbf{v}$ of a particle initially make an angle θ_0 with the direction of the magnetic induction $\mathbf{B}$. Then, since the magnetic moment is constant, we have the equation (Fig. 49)

$$\frac{\sin^2\theta_0}{B_0} = \frac{\sin^2\theta}{B} \tag{37.29}$$

Fig. 49



The reflection of the particle occurs at the point where $\sin\theta = 1$. At this point, the magnetic field must have the value

$$B = \frac{B_0}{\sin^2\theta_0} \qquad (37.30$$

At such a point all particles which, initially, had velocity vectors lying outside a cone with a vertex angle θ_0 will be reflected. Thus, the condition of reflection is independent of the absolute magnitude of the velocity of the particle. All particles with velocity vectors initially inside the cone of angle θ_0 will not be reflected at the point defined by (37.30). The phenomenon of reflection from a magnetic mirror is used for making apparatus used to confine particles to a finite region of space for a long time. Due to the same phenomenon, there are natural radiation belts close to the earth. Charged particles move in the magnetic field of the earth, revolving about the lines of force, and moving along them along the meridians (Fig. 50). Close to the poles, the magnetic field grows stronger and the lines of force

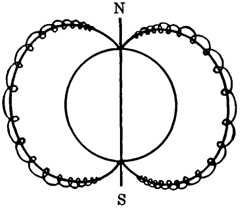


Fig. 50

converge. The particles are, therefore, reflected from the regions close to the magnetic poles, and once again move along the meridians. Since the magnetic field of the earth is inhomogeneous, growing weaker as the distance from the earth increases, there is, in addition to the motion along the meridians, drift in a direction perpendicular to the meridians, i.e., there is a change in longitude, and the particles gradually move along all possible meridians.

Motion in a Transverse Electric Field. Let us now consider the case of a particle performing a motion that is only slightly nonrectilinear, i.e., the radius of curvature of its trajectory is much greater than the region of motion. Let the electric field be in the x direction, and let there be no magnetic field (Fig. 51)

$$E_x = E \qquad E_y = E_z = 0 \qquad \mathbf{B} = 0$$

The equations of motion and the initial conditions are of the form

$$m\frac{d^2x}{dt^2} = eE(z) \qquad m\frac{d^2y}{dt^2} = 0 \qquad m\frac{d^2z}{dt^2} = 0$$
$$x(0) = 0 \qquad \frac{dx(0)}{dt} = 0 \tag{37.31}$$

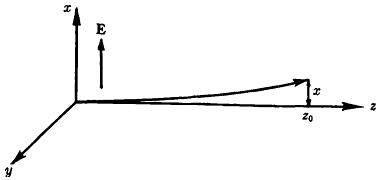


Fig. 51

The field E may vary arbitrarily in the z direction, but its main component must be along the x direction. The trajectory of the particle lies in the xz plane. The x component of the velocity of the particle is much less than the z component, since, by definition, the particle is only slightly deflected from rectilinear motion

$$\frac{v_x}{v_z} = \tan \alpha \ll 1 \tag{37.32}$$

The velocity v of the particle may be written in the form

$$v = \sqrt{v_x^2 + v_z^2} = v_z \left(1 + \frac{v_x^2}{v_z^2}\right)^{1/2} \approx v_z + \frac{1}{2} v_z \frac{v_x^2}{v_z^2} + \cdots$$

Hence, with accuracy of the order of $v_x/v_z \ll 1$, we may assume that

$$v_z = v = \text{const} \tag{37.33}$$

In this approximation, we have

$$\frac{dx}{dt} = \frac{dx}{dz}\frac{dz}{dt} = \frac{dx}{dz} v \qquad \frac{d^2x}{dt^2} = \frac{d^2x}{dz^2} v^2 \tag{37.34}$$

Hence, the equation of motion and the initial conditions (37.31) take the form

$$\frac{d^2x}{dz^2} = \frac{e}{mv^2} E(z) \qquad \left.\frac{dx}{dz}\right|_{z=0} = 0 \qquad \left. x \right|_{z=0} = 0 \tag{37.35}$$

The solution of this equation is obtained with the aid of two quadratures

$$x(z_0) = \frac{e}{mv^2} a \tag{37.36}$$

where

$$a = \int_0^{z_0} d\xi \int_0^{\xi} E(\eta)\, d\eta = \int_0^{z_0} (z_0 - \eta) E(\eta)\, d\eta \tag{37.37}$$

depends only on the configuration of the electric field and is called the *instrumental constant* of the apparatus, in which the electric field under consideration is set up.

Motion in a Transverse Magnetic Field. Let us consider the motion of a particle in a transverse magnetic field (Fig. 52) under the conditions of small deflection

$$B_x = B(z) \qquad B_y = B_z = 0 \qquad \mathbf{E} = 0 \tag{37.38}$$

The magnetic field may vary arbitrarily in the z direction, but its main component must always act in the x direction. Ignoring terms proportional to v_y compared with terms proportional to v_z (since $v_y/v_z \ll 1$), we may write the equation of motion in the field given by (37.38)

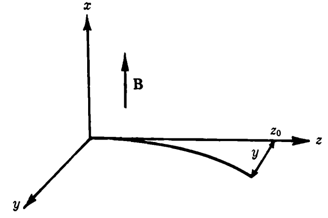


Fig. 52

$$m\frac{d^2x}{dt^2} = 0 \qquad m\frac{d^2y}{dt^2} = evB \qquad m\frac{d^2z}{dt^2} = 0 \tag{37.39}$$

The solution of these equations is obtained as in the preceding case. We thus obtain the equation

$$\begin{gathered} y(z_0) = \frac{e}{mv}b \\ b = \int_0^{z_0} d\eta \int_0^{\eta} B(\xi)\,d\xi = \int_0^{z_0} (z_0 - \eta)B(\eta)\,d\eta \end{gathered} \tag{37.40}$$

for the deflection where b is the instrumental constant of the apparatus, governed by the actual configuration of the field in the apparatus.

Determination of e/m for a Charged Particle. If a particle moves in transverse electric and magnetic fields both acting along the x axis, then the electric field deflects it in the x direction, and the magnetic field in the y direction. Knowing the deflections x and y, defined by (37.36) and (37.40), we can determine the value of e/m of the particle. Squaring both sides of (37.40), and dividing term-by-term by equation (37.36), we obtain

$$\frac{y^2}{x} = \frac{e}{m}\frac{b^2}{a} \tag{37.41}$$

This equation gives e/m, since all the other quantities in it are known. If we know the charge from other measurements (it can only be a multiple of the elementary charge), we can also find the mass of the particle. This method of determining the mass of particles (atoms, molecules, nuclei, etc.) from their motion in electromagnetic fields is widely used in practice in mass spectrometers and mass spectrographs. In these devices, the electromagnetic field is chosen in such a way that particles with different values of e/m move along different trajectories, and the aim is to produce a large difference between the trajectories for a small difference in the value

of e/m. This principle is the basis of the electromagnetic separation of isotopes. As is well known, isotopes consist of nuclei which have the same number of protons but different numbers of neutrons. Consequently, different isotopes have the same charge but different masses, i.e., different values of e/m. In the fields under discussion, different isotopes move along different trajectories, and so become separated. In principle, any mass spectrograph or spectrometer can be used for the separation of isotopes, but in practice apparatus specially constructed for the purpose is employed.

Using such apparatus, it is possible to measure the masses of atoms and elementary particles. These methods are widely used.

§38. Radiation of an Oscillating Electron. Radiation Damping

Free Oscillations of an Elastically Bound Electron. Let the restoring force acting on an electron, when it is displaced from its position of equilibrium, be proportional to the displacement. Taking the origin at the point of equilibrium, and the z axis in the direction of displacement of the electron from the position of equilibrium, we may write the equation of motion of the electron in the form

$$m\ddot{z} + kz = 0 \tag{38.1}$$

The solution of this equation is written in the form

$$z = a \sin \omega t + b \cos \omega t \tag{38.2}$$

where $\omega^2 = k/m$, and a and b are arbitrary constants. The energy of the oscillating electron is equal to

$$W = \frac{m\dot{z}^2}{2} + \frac{m\omega^2}{2} z^2 = \frac{m\omega^2}{2}(a^2 + b^2) \tag{38.3}$$

Radiation of an Oscillating Electron. Imagine a positive charge at the origin, equal in magnitude to the electron charge. This charge remains stationary, and, by Coulomb's law, sets up an electric field, constant in time, which is inversely proportional to the square of the distance. The combination of the moving electron and the stationary positive charge constitute a dipole, the moment of which varies with time. The radiation of such a dipole has been considered in §29. The electromagnetic field vectors of the radiation are inversely proportional to the first power of the distance. It is clear that there is no relationship between the radiation of this electromagnetic field and the constant field of the stationary positive charge, which is inversely proportional to the square of the distance. The radiation field is due to the oscillation of the electron, i.e., it is the radiation field of the oscillating electron. The positive charge which we imagined to

be placed at the origin was only introduced formally so that we might use the equations of §29.

When the electron is at a distance $z(t)$ from the origin, the dipole moment is

$$\mathbf{p}(t) = -|e|z(t)\mathbf{k} \tag{38.4}$$

where $\mathbf{k}$ is a unit vector in the z direction. The minus sign is due to the fact that the direction of the dipole is a vector pointing from the negative to the positive direction. The constants a and b in equation (38.2) are defined by the initial conditions. We can always choose the time origin in such a way that a or b becomes zero. Hence, choosing the time origin in a suitable manner, the harmonic oscillations (38.2) may be written in the form

$$z = b \cos \omega t \tag{38.5}$$

Substituting this expression for z in (38.4), we obtain

$$\mathbf{p} = -\mathbf{k}|e|b \cos \omega t \tag{38.6}$$

Comparison of (38.6) with the real part of expression (29.27a) for a dipole shows that the vector $\mathbf{p}_0$ in (29.27a) is related to the quantities which describe the motion of the electron by the equation

$$\mathbf{p}_0 = -\mathbf{k}|e|b \qquad |\mathbf{p}_0| = |e|b \tag{38.7}$$

Equation (29.33), which describes the radiation field vectors, takes the form

$$\begin{gathered} E_\alpha = E_r = 0 \qquad B_r = B_\theta = 0 \\ cB_\alpha = E_\theta = -\frac{\omega^2}{4\pi\epsilon_0 c^2}\frac{\sin\theta}{r}|e|b\cos\omega\left(\tau - \frac{r}{c}\right) \end{gathered} \tag{38.8}$$

where the time for the wave to reach a sphere of radius r is denoted, for convenience, by τ. From (38.5) it follows that

$$\ddot{z} = -\omega^2 b \cos \omega\tau \tag{38.9}$$

and, hence, (38.8) may be rewritten

$$cB_\alpha = E_\theta = \frac{|e|}{4\pi\epsilon_0 c^2}\frac{\sin\theta}{r}\frac{\partial^2 z\left(\tau - \frac{r}{c}\right)}{\partial\tau^2} = -\frac{e}{4\pi\epsilon_0 c^2}\frac{\sin\theta}{r}\frac{\partial^2 z\left(\tau - \frac{r}{c}\right)}{\partial\tau^2} \tag{38.10}$$

where the fact that the electron charge is negative ($e = -|e|$) is taken into account.

Equation (29.36) for the energy flow through the surface of a sphere is written, using (38.9), in the form

$$Q = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\left\{\frac{\partial^2 z\left(\tau - \frac{r}{c}\right)}{\partial\tau^2}\right\}^2 = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\ddot{z}^2 \tag{38.11}$$

where

$$\ddot{z} = \frac{d^2z(t)}{dt^2} \qquad t = \tau - \frac{r}{c}$$

The mean energy radiated by the oscillating electron in one period is determined from the equation

$$\langle Q\rangle = \frac{1}{12\pi\epsilon_0}\frac{e^2}{c^3}\omega^4 b^2 \tag{38.12}$$

This is the mean total radiation flux cutting a sphere of radius r at time τ. It is clear that this energy was emitted by the oscillating electron during the preceding instants of time.

Hence, equation (38.12) describes the rate of loss of the energy of the electron W by radiation. By the law of conservation of energy, we may write

$$\frac{dW}{dt} = -\langle Q\rangle = -\frac{1}{12\pi\epsilon_0}\frac{e^2}{c^3}\omega^4 b^2 \tag{38.13}$$

From equation (38.3) (for $\alpha = 0$), it follows that

$$b^2 = \frac{2}{m\omega^2}W$$

Hence, equation (38.13) may be rewritten

$$\frac{dW}{dt} = -\gamma W \qquad \gamma = \frac{1}{6\pi\epsilon_0}\frac{e^2\omega^2}{mc^3} \tag{38.14}$$

The solution of this equation is of the form

$$W(t) = W_0 e^{-\gamma t} \tag{38.15}$$

where W_0 is the energy of the oscillating electron at time $t = 0$.

From equation (38.3), which relates the energy of the oscillating electron to the amplitude of the oscillation, it is evident that the reduction in the energy of the electron takes place according to the law (38.15) in the case if a and b vary according to the law

$$a = a_0 e^{-\frac{\gamma}{2}t} \qquad b = b_0 e^{-\frac{\gamma}{2}} \tag{38.16}$$

Thus, allowing for the loss of energy due to radiation, instead of equation (38.2), we must write

$$z = e^{-\frac{\gamma}{2}t}(a_0 \sin\omega t + b_0 \cos\omega t) \tag{38.17}$$

Consequently, when the radiation is taken into account, the equation of motion of the electron (38.1) must be changed. It should be supplemented by an additional force which describes the effect of the radiation damping on the electron.

Force on the Electron. The radiation force on the electron may be represented physically as the reaction of the radiation field. As we shall show later, electromagnetic waves possess momentum. By the law of conservation of momentum, the momentum of a closed system (electron + radiation) must be constant. Hence, the momentum of the electron must change by the value of the momentum of the electromagnetic wave which it radiates. This is equivalent to saying that during radiation, a force acts on the electron. Since the energy and velocity of the electron decrease as a result of radiation, it follows that this is a decelerating (damping) force.

Supplementing equation (38.1) by a force F which describes the damping due to radiation we have

$$m\ddot{z} + kz = F \tag{38.18}$$

Multiplying both sides of this equation by $\dot{z}$

$$\frac{d}{dt}\left(\frac{m\dot{z}^2}{2} + \frac{k}{2}z^2\right) = F\dot{z} \tag{38.19}$$

On the right-hand side, we have the work done by the damping force per unit time. By definition, this is equal to the radiation power, defined by (38.11)

$$F\dot{z} = -\frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\ddot{z}^2 \tag{38.20}$$

This equation expresses the law of conservation of energy. In general, it is not possible to express F as a linear function of z and its derivatives. This may be done only approximately, assuming that: (a) the damping of the oscillations is not very strong, so that for a fairly small number of periods, the motion may be considered periodic; (b) it is sufficient to formulate an average law of conservation of energy for a fairly small number of periods.

We average with respect to time the obvious equation

$$\ddot{z}^2 = -(\dddot{z}\,\dot{z}) + \frac{d}{dt}(\ddot{z}\,\dot{z})$$

using only the stated assumptions. Since the motion is periodic, we have

$$\left\langle\frac{d}{dt}(\dot{z}\,\ddot{z})\right\rangle = \frac{1}{T}\{(\dot{z}\,\ddot{z})_{t=T} - (\dot{z}\,\ddot{z})_{t=0}\} = 0$$

and, hence, equation (38.20) may be written

$$\langle F\dot{z}\rangle = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\langle(\dddot{z}\ \dot{z})\rangle$$

The above equation will be satisfied if we put

$$F = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\dddot{z} \tag{38.21}$$

Thus, the equation of motion (38.18), including the damping force due to radiation, has the form

$$m\ddot{z} + kz - \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\dddot{z} = 0 \tag{38.22}$$

The solution of this equation should have the form of (38.17). Let us seek it in complex form

$$z = de^{i\omega_\gamma t} \tag{38.23}$$

Substitution of this equation in (38.22) gives the equation defining ω_γ

$$-\omega_\gamma{}^2 + \omega^2 + i\gamma\omega^{-2}\omega_\gamma{}^3 = 0 \tag{38.24}$$

where

$$\omega = \frac{k}{m} \qquad \gamma = \frac{1}{6\pi\epsilon_0}\frac{e^2\omega^2}{mc^3}$$

By definition, the damping is weak. This means that for an interval of time of the order of the oscillation period $T = 2\pi/\omega$, the exponential factor in (38.16) is only slightly different from unity, i.e., the inequality

$$\gamma \ll \omega \tag{38.25}$$

holds. For $\gamma = 0$, the solution of (38.24) takes the form

$$\omega_\gamma = \pm\omega \tag{38.26}$$

For small values of γ, this solution must be sought in the form

$$\omega_\gamma = \pm\omega + \epsilon \tag{38.27}$$

where $\epsilon \ll \omega$. Substituting (38.27) in (38.24), and ignoring terms of the order of ϵ^2, $\gamma\epsilon$, and higher, we obtain

$$\epsilon = i\frac{\gamma}{2}$$

Thus, for ω_γ we obtain the expression

$$\omega_\gamma = \pm\omega + i\frac{\gamma}{2} \tag{38.28}$$

and we use it to obtain the general solution (38.23) for z in the form

$$z = e^{-\frac{\gamma}{2}t}(d_1 e^{i\omega t} + d_2 e^{-i\omega t}) \tag{38.29}$$

The real and imaginary parts of this expression have the form of (38.17).

Expression (38.21) for the force is applicable only to almost periodic motion when the damping is sufficiently weak. In this case it may be rewritten in the form

$$F = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\dddot{z} \approx -\frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\omega^2\dot{z} \tag{38.30}$$

which shows that the force acts opposite to the velocity $\dot{z}$.

Taking (38.30) into account, the equation of motion (38.22) may be written in the form

$$\ddot{z} + \gamma\dot{z} + \omega^2 z = 0 \tag{38.31}$$

Generalized Radiation Formula. Equation (38.11) states that the instantaneous radiated power at some instant of time is governed by the acceleration of the charged particle at that instant. Consequently, the power radiated at a given instant is independent of how the particle is moving before that instant and how it will move after that instant. Hence, although equation (38.11) has been deduced for the case of harmonic motion, it is true for any motion, provided that $\ddot{z}^2$ is taken to mean the square of the acceleration of the particle in the motion under discussion. Hence, we conclude that if in some arbitrary motion the position of the particle is given by the radius vector $\mathbf{r}(t)$, then the power radiated by that particle is equal to

$$Q = \frac{dW}{dt} = -\frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}(\ddot{\mathbf{r}})^2 \tag{38.32}$$

This expression is true for velocities that are small in comparison with the velocity of light. The generalization of this equation to the case of relativistic velocities will be given in Part III.

§39. Theory of the Spectral Line Width

If an electron oscillates without damping, and radiates for an infinite interval of time, then the wave it radiates is strictly monochromatic, i.e., its wavelength is strictly defined. However, in fact, an electron radiates for a finite interval of time. Even if during this finite interval, the electron oscillates with a strictly defined frequency, the wave which it radiates cannot be monochromatic, because if this wave is expanded as a Fourier

integral, then all possible harmonics will occur in it, i.e., waves radiated during a finite interval of time cannot be strictly monochromatic. If during the process of radiation, the amplitude of the oscillations of the electron decreases, then the wave it radiates will not be monochromatic either. This case, in essence, reduces to the preceding case, since after some finite interval of time the oscillations practically cease, due to damping, and, consequently, the radiation stops too, i.e., we have a wave radiated for a finite interval of time.

Since the wave is not monochromatic, the energy radiated is distributed over all possible frequencies. If the duration of the radiation is infinitely long, then all the radiation energy has strictly one frequency. A finite duration of radiation leads to the radiation energy being distributed over some closely spaced band of frequencies, and the band width increasing as the duration of radiation becomes shorter. If this radiation is observed with a spectroscope, then we shall see a radiation line of finite width, consisting of all lines of the various frequencies. A strictly monochromatic wave would give an infinitely narrow line. Since the duration of radiation is finite, the line is of finite width. The factors which decrease the duration of radiation, e.g., the damping of the oscillations, and also collisions between molecules (since collisions affect the duration of radiation of molecules), must lead to an increase in the width of the spectral lines. Let us now consider the increase in width of spectral lines caused by the damping of oscillations.

If an electron begins to oscillate at time $t = 0$, then its displacement $z(t)$ from the position of equilibrium, taking the damping into account, is given by

$$z(t) = \begin{cases} 0 & \text{for } t < 0 \\ e^{-\frac{\gamma}{2}t}(ae^{i\omega_0 t} + a^* e^{-i\omega_0 t}) & \text{for } t > 0 \end{cases} \tag{39.1}$$

where a^* is the complex conjugate of a, so that the value of $z(t)$ is real. The frequency of free oscillations of the electron will be denoted, for convenience, by ω_0.

The energy W radiated in the interval $0 < t < \infty$ is defined, on the basis of (38.11), by

$$W = \int_0^\infty Q\,dt = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\int_{-\infty}^{\infty} \ddot{z}^2\,dt \tag{39.2}$$

since $\ddot{z} = 0$ when $t < 0$.

The problem is to expand (39.2) as a Fourier integral, i.e., to express it in the form

$$W = \int_0^\infty W_\omega \, d\omega \tag{39.3}$$

Since the damping is weak, we may assume that

$$\ddot{z} \approx \omega_0^2 z \tag{39.4}$$

We now express z, defined by (39.1), in terms of a double Fourier integral

$$z(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \, d\omega \int_{-\infty}^{\infty} z(\xi) e^{i\omega\xi} \, d\xi = \int_{-\infty}^{\infty} z_\omega e^{-i\omega t} \, d\omega \tag{39.5}$$

introducing the notation

$$z_\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} z(\xi) e^{i\omega\xi} \, d\xi = \frac{1}{2\pi} \int_0^{\infty} z(\xi) e^{i\omega\xi} \, d\xi \tag{39.5a}$$

since $z = 0$ for $\xi < 0$. Using (39.4) and (39.5), we can put (39.2) in the form

$$W = \frac{1}{6\pi\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} z_\omega e^{-i\omega t} \, d\omega \int_{-\infty}^{\infty} z_\xi e^{-i\xi t} \, d\xi$$

$$= \frac{1}{6\pi\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \iint_{-\infty}^{\infty} z_\omega z_\xi \, d\omega \, d\xi \int_{-\infty}^{\infty} e^{-i(\omega+\xi)t} \, dt \tag{39.6}$$

But

$$\int_{-\infty}^{\infty} e^{-i(\omega+\xi)t} \, dt = 2\pi \, \delta(\omega + \xi) \tag{39.7}$$

where $\delta(\omega + \xi)$ is a δ-function. Hence, (39.6) becomes

$$W = \frac{1}{3\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \iint_{-\infty}^{\infty} z_\omega z_\xi \, d\omega \, d\xi \, \delta(\omega + \xi) = \frac{1}{3\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \int_{-\infty}^{\infty} z_\omega z_{-\omega} \, d\omega \tag{39.8}$$

where we have integrated with respect to ξ. Since

$$f(\omega) = z_\omega z_{-\omega} \tag{39.9}$$

is an even function of ω, the integral in (39.8) may be written

$$\int_{-\infty} z_\omega z_{-\omega} \, d\omega = 2 \int_0^\infty z_\omega z_{-\omega} \, d\omega$$

Hence, finally, we have

$$W = \int_0^\infty \frac{2}{3\epsilon_0} \frac{e^2}{c^3} \omega_0^4 z_\omega z_{-\omega} \, d\omega \equiv \int_0^\infty W_\omega \, d\omega \tag{39.10}$$

The quantity z_ω is defined by (39.5a), where $z(\xi)$ has the form (39.1). Let us evaluate it

$$z_\omega = \frac{1}{2\pi}\int_0^\infty \left\{ae^{\left[-\frac{\gamma}{2}+i(\omega_0+\omega)\right]\xi} + a^*e^{\left[-\frac{\gamma}{2}-i(\omega_0-\omega)\right]\xi}\right\} d\xi$$

$$= \frac{1}{2\pi}\left\{\frac{a}{\frac{\gamma}{2} - i(\omega_0+\omega)} + \frac{a^*}{\frac{\gamma}{2} + i(\omega_0-\omega)}\right\} \qquad (39.11)$$

In its general form, the expression for $z_\omega z_{-\omega}$ is very cumbersome. However, we know that when the damping is weak ($\gamma \ll \omega_0$), the spectral lines are narrow, and, consequently, we are only interested in the frequencies close to ω_0. Hence, we may assume that $|\omega - \omega_0| \ll \omega_0$. Under these conditions, the term in (39.11) which contains $(\omega_0 - \omega)$ in the denominator is much greater in absolute value than the term containing $(\omega_0 + \omega)$ in the denominator. Hence, to find $z_\omega z_{-\omega}$, we keep only the principal terms, which have $(\omega_0 - \omega)$ in the denominator. We thus obtain

$$z_\omega z_{-\omega} = \frac{1}{2\pi}\frac{a^*}{\frac{\gamma}{2} + i(\omega_0-\omega)}\frac{1}{2\pi}\frac{a}{\frac{\gamma}{2} - i(\omega_0-\omega)} = \frac{1}{4\pi^2}\frac{a^*a}{\frac{\gamma^2}{4} + (\omega_0-\omega)^2} \qquad (39.12)$$

Hence, the value of W_ω which describes the form of the spectral lines is given by

$$W_\omega = \frac{1}{6\pi^2\epsilon_0}\frac{e^2}{c^3}\omega_0^4\frac{a^*a}{\frac{\gamma^2}{4} + (\omega_0-\omega)^2} \qquad (39.13)$$

This expression states that the maximum intensity in the radiation spectrum occurs at the frequency $\omega = \omega_0$. The intensity decreases rapidly away from this maximum. At $\omega - \omega_0 = \gamma/2$ from the maximum, the intensity falls to half its maximum value (Fig. 53). This half-width of the spectral line is equal to

$$\delta\omega = \gamma \qquad (39.14)$$

The width of a spectral line defines the interval of frequencies close to the fundamental frequency ω_0 within which the main part of the energy is radiated. Equation (39.14) states that the width of a spectral line is inversely proportional to the duration of radiation. If τ_{rad} denotes the time during which the amplitude of oscillation of an electron decreases by a factor e, then equation (39.14) may be written

$$\delta\omega = \frac{2}{\tau_{\text{rad}}} \qquad (39.15)$$

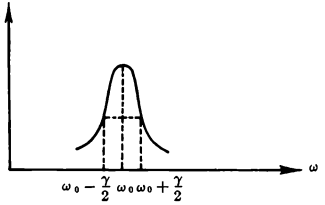


Fig. 53

§40. Scattering of Light by Free Electrons

Let a plane-polarized wave fall upon a free electron from the y direction. The electric vector of the wave lies in the z direction. Under the action of the electric field of the wave, the electron begins to move, the motion being given by

$$m_0\ddot{z} = eE(t) \tag{40.1}$$

During its motion, the electron will, in general, also experience a force due to the electromagnetic field of the wave, and this force bears approximately the same ratio to the electric force as does the velocity of the electron to the velocity of light. Thus, for nonrelativistic velocities of the electron, the magnetic field may be ignored.

During its motion described by equation (40.1), the electron radiates an electromagnetic wave, the power radiated, from (38.11), being

$$Q = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\ddot{z}^2 = \frac{1}{6\pi\epsilon_0}\frac{e^2}{c^3}\frac{e^2E^2}{m_0^2} \tag{40.2}$$

The direction of motion of the electromagnetic waves radiated by the electron is different from the direction of motion of the waves falling on the electron. Thus, the whole process of the radiation of an electromagnetic wave by an electron set in motion by another electromagnetic wave may be considered as the scattering of incident electromagnetic waves by a free electron. The intensity of scattering is described by the effective scattering cross section, which is defined as the area which the cross section of the electron must have for the incident-wave energy reaching this area to be equal to the energy scattered by the electron.

The density of the energy flux in the incident wave is defined by the Poynting vector

$$P = \sqrt{\frac{\epsilon_0}{\mu_0}} E^2 \tag{40.3}$$

This is the energy falling in unit time, on unit area, perpendicular to the direction of propagation of the wave. If the effective cross section for the scattering of light by a free electron is denoted by σ, then, by definition, we may write

$$P\sigma = Q \tag{40.4}$$

where Q is defined by (40.2). Hence, we obtain

$$\sigma = \frac{8\pi}{3} r_0^2 \tag{40.5}$$

where

$$r_0 = \frac{e^2}{4\pi\epsilon_0 m_0 c^2} = 2.8 \times 10^{-15}\,\mathrm{m} = 2.8 \times 10^{-13}\,\mathrm{cm}$$

is called the *classical radius* of the electron. Thus, if the electron is represented as a spherical particle, which scatters electromagnetic waves when they strike its surface, then the radius of the electron must be taken equal to $\sqrt{\frac{8}{3}}r_0$. In fact, the mechanism of scattering described in this paragraph assumes that the electron has the dimensions of a point, and the assumption that it has finite dimensions is only introduced as a means of interpreting the effective cross section which describes the intensity of scattering.

§41. Momentum of an Electromagnetic Field. Pressure of Light

Let us consider some volume V, in which there is an electromagnetic field interacting with charges enclosed by this volume. The electromagnetic field exerts a force on this volume equal to

$$\mathbf{F} = \int_V \mathbf{f}\, dV \tag{41.1}$$

where $\mathbf{f} = \rho\mathbf{E} + \rho\mathbf{v} \times \mathbf{B}$ is the density of the Lorentz force.

Using Maxwell's equations in the form

$$\rho\mathbf{v} = \operatorname{curl}\mathbf{H} - \frac{\partial\mathbf{D}}{\partial t} \qquad \rho = \operatorname{div}\mathbf{D} \tag{41.2}$$

we can eliminate ρ and $\mathbf{v}$ from the expression for the density of the Lorentz force. We thus obtain

$$\mathbf{f} = \mathbf{E}\operatorname{div}\mathbf{D} + \operatorname{curl}\mathbf{H} \times \mathbf{B} - \frac{\partial\mathbf{D}}{\partial t} \times \mathbf{B} \tag{41.3}$$

Since

$$\operatorname{div} \mathbf{B} = 0 \qquad \operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{41.4}$$

equation (41.3) is easily transformed into

$$\mathbf{f} = \mathbf{E} \operatorname{div} \mathbf{D} + \mathbf{H} \operatorname{div} \mathbf{B} + \operatorname{curl} \mathbf{H} \times \mathbf{B} + \operatorname{curl} \mathbf{E} \times \mathbf{D} - \frac{\partial}{\partial t} (\mathbf{D} \times \mathbf{B}) \tag{41.5}$$

We shall consider, for example, the x component of this force

$$f_x = E_x \operatorname{div} \mathbf{D} + H_x \operatorname{div} \mathbf{B} + (\operatorname{curl} \mathbf{H} \times \mathbf{B})_x + (\operatorname{curl} \mathbf{E} \times \mathbf{D})_x - \frac{\partial}{\partial t} (\mathbf{D} \times \mathbf{B})_x$$

It may be verified directly that this expression can be put in the form

$$f_x = \operatorname{div} \mathsf{F} - \frac{\partial}{\partial t} (\mathbf{D} \times \mathbf{B})_x \tag{41.6}$$

where the components of the vector F are defined by

$$\mathsf{F}_x = D_x E_x + H_x B_x - \frac{1}{2} (\mathbf{D} \cdot \mathbf{E} + \mathbf{H} \cdot \mathbf{B})$$

$$\mathsf{F}_y = D_x E_y + H_x B_y$$

$$\mathsf{F}_z = D_x E_z + H_x B_z$$

Furthermore

$$\mathbf{D} \times \mathbf{B} = \frac{1}{c^2} \mathbf{E} \times \mathbf{H}$$

since $\epsilon_0 \mu_0 = 1/c^2$.

Thus, the x component of the force acting on the charges in V (see eq. (41.1)), may be written in the form

$$F_x = \int_V f_x \, dV = \int_V \operatorname{div} \mathsf{F} \, dV - \frac{d}{dt} \int_V \frac{1}{c^2} (\mathbf{E} \times \mathbf{H})_x \, dV \tag{41.7}$$

Similar equations may be obtained for the y and z components.

The force $\mathbf{F}$ acting on the charges in V imparts an acceleration to these charges. We shall denote the total momentum of the charged particles in V by $\mathbf{G}^M$. Then, by Newton's law, we may write

$$\frac{d\mathbf{G}^M}{dt} = \mathbf{F} \tag{41.8}$$

Hence, equation (41.7) takes the form

$$\frac{d}{dt} \left(G_x^M + \int \frac{1}{c^2} (\mathbf{E} \times \mathbf{H})_x \, dV \right) = \oint_S \mathsf{F} \cdot d\mathbf{S} \equiv F_x^{\text{surf}} \tag{41.9}$$

where we have used Gauss' theorem

$$\int_V \operatorname{div} \mathsf{F} \, dV = \oint_S \mathsf{F} \cdot d\mathbf{S} \equiv F_x^{\text{surf}}$$

F_x^{surf} is the x component of the force acting on the surface of the volume under consideration. Formulas similar to (41.9) are also obtained for the y and z components. Hence, passing from components to vectors, instead of (41.9), we may write

$$\frac{d}{dt}(\mathbf{G}^M + \mathbf{G}^{\text{field}}) = \mathbf{F}^{\text{surf}} \tag{41.10}$$

introducing the notation

$$\mathbf{G}^{\text{field}} = \int_V \mathbf{g}\, dV \qquad \mathbf{g} = \frac{1}{c^2}\mathbf{E} \times \mathbf{H} = \frac{\mathbf{P}}{c^2}$$

Equation (41.10) states that if the volume under consideration is an insulated system, and if no forces are applied to its surface $\mathbf{F}_{\text{surf}} = 0$, then

$$\frac{d\mathbf{G}^M}{dt} = -\frac{d\mathbf{G}^{\text{field}}}{dt} \tag{41.11}$$

This means that the momentum of the matter in V changes as a result of the change in $\mathbf{G}^{\text{field}}$. Hence, $\mathbf{G}^{\text{field}}$ is the momentum of the electromagnetic field enclosed in V, and $\mathbf{g}$ is the density of the momentum of the electromagnetic field.

Pressure of Light. If an electromagnetic wave falls on a body and is absorbed by it, then the momentum of the wave is transferred to the body. This means that a force acts on the body which appears as a pressure (pressure of light in the case of light waves).

Let us evaluate this pressure. It is equal to the momentum transferred by the wave to the body per unit time per unit surface. Thus, when the electromagnetic wave is normal to the surface and totally absorbed, the pressure is equal to

$$p = cg = \frac{1}{c}EH = \epsilon_0 E^2 = u \tag{41.12}$$

where $u = (ED + HB)/2 = \epsilon_0 E^2$ is the energy density of the electromagnetic field in the electromagnetic wave. Thus, the pressure of the electromagnetic wave is, in this case, numerically equal to the energy density of this wave.

If the surface of the body is perfectly reflecting, the momentum of the wave is reflected in the opposite direction. In this case, therefore, the momentum transmitted to the body is twice as great as in the case of total absorption. The pressure, therefore, is also twice as great as in the case of complete absorption.

If the electromagnetic wave strikes the surface at some angle, then the calculation of the pressure is carried out in a similar manner, taking the

component of the momentum of the wave perpendicular to the surface. The case of partial reflection and partial absorption is also considered in a similar way.

PROBLEMS

1 In an atom of hydrogen, the electron moves in a circular orbit at a distance $r_1 = 5.3 \times 10^{-9}$ cm from the nucleus. The nucleus (proton) may be taken to be stationary. Determine the force of attraction between the electron and proton, the electric field of the proton at the electron orbit, and the velocity of the electron.

Solution:

$$F = \frac{1}{4\pi\epsilon_0}\frac{e^2}{r_1^2} = 8 \times 10^{-8}\ \mathrm{N}$$

$$E = \frac{1}{4\pi\epsilon_0}\frac{|e|}{r_1^2} \approx 0.58 \times 10^{12}\ \mathrm{V/m} = 0.58 \times 10^{10}\ \mathrm{V/cm}$$

$$\frac{m_0 v^2}{r_1} = |e|E \qquad v = \sqrt{\frac{|e|Er_1}{m_0}} \approx 2.2 \times 10^6\ \mathrm{m/sec}$$

2 An electron moving through transverse homogeneous magnetic and electric fields acting in the same direction, is deflected over a distance $l = 5$ cm from its initial direction by an angle of 7/110 radians with respect to each field. Determine the ratio of the charge of the electron to its mass, and find the velocity of the electron: $B_0 = 3.6$ G, $E_0 = 100$ V/cm.

Solution: Using equations (37.36) and (37.40), we have

$$a = \int_0 (l - \eta)E_0\, d\eta = \frac{l^2}{2}E_0 \qquad b = \int_0 (l - \eta)B_0\, d\eta = \frac{l^2}{2}B_0$$

$$\frac{e}{mv^2}E_0\frac{l^2}{2} = \frac{e}{mv}B_0\frac{l^2}{2}$$

$$v = \frac{E_0}{B_0} \approx 2.8 \times 10^7\ \mathrm{m/sec} \qquad \frac{e}{m} = \frac{\left(l\,\frac{7}{110}\right)v^2 2}{B_0 l^2} \approx 1.76 \times 10^{11}\ \mathrm{coul/kg}$$

3 An electron moves along the lines of force of a homogeneous electric field of intensity $E = 10$ V/cm. The initial velocity of the electron is $v_0 = 10{,}000$ km/sec. Determine the time t and distance x for the electron to come to a complete standstill.

Solution:

$$\frac{m_0 v^2}{2} = |e|Ex \qquad x = \frac{m_0 v^2}{2|e|E} = 0.3\ \mathrm{m} = 30\ \mathrm{cm}$$

$$\frac{|e|E}{m_0}t = v_0 \qquad t = \frac{m_0 v_0}{|e|E} = 5.7 \times 10^{-8}\ \mathrm{sec}$$

4 Determine the total momentum of electrons representing an electron current $I = 200$ amps in a right cylindrical volume of length $l = 5$ km. The electrons move along the axis of the cylinder.
Solution:

$$j = \rho v \qquad \rho = |e|N$$

where N is the number of electrons per unit volume

$$I = jS$$

where S is the cross section of the cylinder

$$p = NSlm_0v = \frac{Ilm_0}{|e|} = 5.7 \times 10^{-6} \text{ kg m/sec}$$

5 Determine the radius of curvature of the trajectory of an electron possessing energy $W = 100$ eV in a magnetic field $B = 10^3$ G.

$$\textit{Answer:} \quad r = \frac{\sqrt{2Wm_0}}{|e|B} = 3.4 \times 10^{-4} \text{ m} = 0.34 \text{ mm}$$

6 Determine the velocity of drift in crossed electric and magnetic fields; $E = 200$ V/cm, $B = 500$ G.

$$\textit{Answer:} \quad v_0 = \frac{E}{B} = 4 \times 10^5 \text{ m/sec}$$

7 A beam of electromagnetic waves, cross section $S = 10^{-2}$ cm² and energy flux density $P = 10$ w/cm², falls on a thin layer of electrons of thickness $d = 10^{-4}$ cm. The electron density in the layer is $N = 10^6$ cm^{-3}. Determine the energy dissipated by the electrons per second.
Hint: Use the value of the scattering cross section of electromagnetic waves incident on a single electron.

$$\textit{Answer:} \quad x = PNS\,d\sigma = 6.5 \times 10^{-24} \text{ W}$$

8 In the earth's orbit, the solar energy flux is approximately equal to 4 cal cm^{-2} min^{-1}. Supposing that 1% of the energy falling on the earth is absorbed, determine the force exerted by the solar radiation on the earth. Take the radius of the earth equal to 6400 km.
Solution:
1 cal/sec = 4.18 W.
Pressure

$$p = u10^{-2} = \frac{P}{c}10^{-2} = 9.4 \times 10^{-8} \text{ N/m}^2$$

Force

$$F = \pi r^2 p = 1.2 \times 10^7 \text{ N}$$

9 Determine the force exerted by an electromagnetic flux of density P on a perfectly reflecting sphere of radius r (Fig. 54).
Solution: The only nonzero component of the force lies along the direction

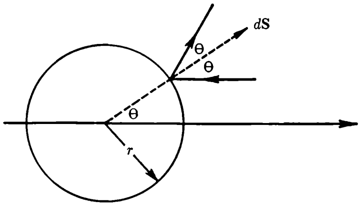


Fig. 54

of the initial flux. The component of the force from the element $d\sigma$ in the direction of the initial flux equals (Fig. 54)

$$dF = 2\frac{P}{c}\cos^2\theta\, d\sigma$$

The cosine occurs once in this formula due to the angle of incidence of the beam on the corresponding element of the sphere, and a second time due to the projection of the force along the direction of the initial flow. The surface element in a spherical system of coordinates with the z axis in the opposite direction to the flux equals

$$d\sigma = r^2 \sin\theta\, d\alpha\, d\theta$$

Hence, we have

$$F = 2\frac{P}{c} r^2 \int_0^{2\pi} d\alpha \int_0^{\frac{\pi}{2}} \cos^2\theta \sin\theta\, d\theta = \frac{4\pi}{3}\frac{P}{c} r^2$$

i.e., in this case the force is $\frac{4}{3}$ times greater than in the case when the whole beam is absorbed by the sphere.

10 What is the pressure of the solar radiation on bodies on the surface of the sun in the case when the rays are completely absorbed by the bodies. The radius of the sun is 700,000 km, and the distance from the sun to the earth is 140×10^6 km.
Hint: Use the data of Problem 8 of this chapter.

Answer: $p = 9.4 \times 10^{-6}\left(\frac{140 \times 10^6}{7 \times 10^5}\right)^2 \approx 0.38\ \text{N/m}^2$

11 Find the radius of an absolutely black spherical particle of density $\rho_m = 5$ gm/cm^3, for which the light pressure due to the sun's radiation is equal to the sun's gravitational attraction. The mass of the sun is 2×10^{30} kg, and the gravitational constant $G = 6.7 \times 10^{-11}$ N m^2/kg^2.
Solution: Since the pressure and the gravitational attraction are inversely proportional to the square of the distance, the solution of the problem is

independent of the distance. It is convenient to write down the expression for the distance from the sun to the earth. Hence, we have

$$\pi r^2 p = G\frac{Mm}{R^2}$$

where r is the unknown radius of the particle, p is the pressure of the sun's rays at the orbit of the earth, R is the distance between the sun and the earth, M is the mass of the sun, $m = \frac{4\pi}{3} r^3\rho_m$ is the mass of the particle. Hence, we find

$$r = \frac{3pR^2}{4GM\rho_m} \approx 10^{-7}\ \mathrm{m} = 10^{-5}\ \mathrm{cm}$$

12 In a homogeneous magnetic field $B = 500$ G, an electron moves in a circle, radius $r = 10$ cm. The field then slowly increases to $B = 4500$ G. Find the radius of the circle along which the electron moves at the end of the process. *Solution:* As the magnetic field increases, an induced electric field acts on the electron, and, hence, the energy of the electron changes. We have to use the adiabatic invariant

$$\frac{\frac{1}{2} m_0 v_\perp^2}{B} = \mathrm{const}$$

Since, on the other hand

$$\frac{m_0 v_\perp^2}{r} = |e| v_\perp B,$$

the condition of conservation of the adiabatic invariant may be written

$$Br^2 = \mathrm{const}$$

Hence, it follows that

$$B_0 r_0^2 = B_1 r_1^2,\ r_1 = r_0\sqrt{\frac{B_0}{B_1}} \approx \frac{10}{3}\ \mathrm{cm}$$

We may consider the problem in another manner. As has been observed, the condition of conservation of the adiabatic invariant is equivalent to saying that the orbit of the electron encloses a magnetic flux of constant magnitude. Hence, we can immediately write down the constant magnetic flux enclosed by the orbit of the electron

$$B_0 \pi r_0^2 = B_1 \pi r_1^2$$

which agrees with the condition previously stated.

13 A beam of protons moving at a speed $v_0 = 3 \times 10^6$ m/sec strikes a thin sheet of foil which slows down, but does not absorb, the protons. Some of the kinetic energy of the protons is transmitted to the foil. For simplicity, it may be assumed that, after passing through the foil, the beam moves in the same direction as before. Determine the pressure which the beam exerts on the foil if the density of the protons in the beam after passing through the foil is doubled and becomes $N_1 = 10^7\ \mathrm{cm}^{-3}$.

Solution: Since the protons are not retained in the foil, and the process is steady, the equation of continuity takes the form

$$\operatorname{div} \mathbf{j} = 0$$

For a parallel beam of protons passing through the foil, this means that the current density of the protons does not change on passing through the foil, i.e., $j_0 = j_1$, where j_0 is the current density before passing through the foil, and j_1 is the current density afterwards. Hence, it follows that

$$\rho_1 v_1 = \rho_0 v_0$$

Consequently

$$v_1 = \frac{\rho_0}{\rho_1} v_0 = \frac{1}{2} v_0$$

After passing through the foil, the current density remains the same since the velocity is decreased by the same factor by which the density of protons in the beam is increased. The pressure is equal to the momentum transmitted per unit time to unit area of the foil

$$p = v_0 N_0 m_0 (v_0 - v_1) = v_1 N_1 m_0 (v_0 - v_1)$$

$$= \frac{v_0^2}{2} N_0 m_0 = \frac{v_0^2}{4} N_1 m_0 = 3.7 \times 10^{-2}\ \mathrm{N/m^2}$$

14 An electron of energy 25 eV begins to move parallel to an earthed conducting plate at a distance of 1 mm from it. What is the distance l which the electron will travel along the plate before it collides with it?

Answer: $l = 13.1$ m

15 A magnetron consists of a straight filament and a concentric cylindrical surface to which a potential difference is applied. A homogeneous magnetic field is applied parallel to the axis of the magnetron. Consider the motion of electrons emitted from the filament and accelerated by the potential difference between the filament and the cylindrical electrode. What is the value of the magnetic field if there is no current between the filament and the cylindrical electrode for a potential difference of less than $V = 8.55$ V.

Answer: $B = 6.66$ G

16 A cylindrical beam of electrons, of radius r, is created by some method. The energy of each electron in the beam is 1 eV. The moving beam of electrons creates a current I. As a result of the interaction between the electrons, the beam expands. Find the acceleration of the peripheral electrons in the beam.

Answer: $\dfrac{d^2r}{dt^2} = \dfrac{I}{2\pi\epsilon_0}\left(\dfrac{e}{2m_0V}\right)^{1/2}\dfrac{1}{r}\left(1 - \dfrac{v^2}{c^2}\right)$

CHAPTER 8

Dielectrics

§42. Rarefied Gases

Polar and Nonpolar Molecules. The electrical properties of a system whose total charge is equal to zero is described, in the first approximation, by the *dipole moment*, defined by the formula

$$\mathbf{p} = \int_V \rho \mathbf{r} \, dV \tag{42.1}$$

The molecules of a substance, taken as a whole, are electrically neutral, and, hence, in the first approximation, the properties of molecules are described by their dipole moments. If, in the absence of an external electric field, the dipole moment of a molecule is zero, the molecule is said to be *nonpolar*. A molecule is said to be *polar* if its dipole moment is nonzero, in the absence of an external electric field.

Molecular Picture of the Polarization of a Dielectric. A volume ΔV is said to be a physically small volume if, on the one hand, it is sufficiently small for variations in macroscopic physical quantities (e.g., **E**, **H**) within it to be ignored, and, on the other hand, it is large enough to contain a large number of molecules. The polarization vector **P** is defined as the dipole moment per unit volume

$$\mathbf{P} = \frac{1}{\Delta V} \sum_i \mathbf{p}_i \tag{42.2}$$

The summation is carried out over the dipole moments $\mathbf{p}_i$ of the molecules in the physically small volume ΔV.

In the absence of an external electric field, the dipole moments of nonpolar molecules are zero, and those of polar molecules are oriented in

space in a random manner. Hence, in both cases, in the absence of an external electric field, a dielectric is not polarized; its polarization vector equals zero.

When a dielectric is introduced into an external electric field, polarization occurs. In the case of nonpolar molecules, this is due to dipole moments induced in the molecules, while in the case of polar molecules it is due to the reorientation of the dipole moments of the molecules along the direction of the electric field. The nature of the polarization in both cases is determined by the magnitude of the electric field acting on the molecules. This field is made up of the external field and of the field set up by all the other molecules of the polarized dielectric. Hence, generally speaking, the field acting on a molecule of a dielectric is not equal to the external field. This is particularly important in the case of dense gases and condensed media. In the case of sufficiently rarefied gases, however, it may be assumed that the field acting on the molecules is equal to the external field. We shall consider this case now, and the case of dense gases and condensed systems in the following section.

It must be pointed out that dipole moments may be induced also in polar molecules. However, this additional dipole moment is generally very small in comparison with the permanent dipole moment of a polar molecule, and, hence, it may be ignored.

Theory of the Polarization of Dielectrics Consisting of Molecules. The electric field forces act on the unlike charges of a molecule in opposite directions: positive charges tend to move in the direction of the field, and negative charges against the field. Consequently, the molecule becomes deformed, and a dipole moment is induced in it by the electric field. From the mechanism of formation of the induced dipole moment, it follows that its direction coincides with the direction of the electric field. In the first approximation, the value of the dipole moment may be taken to be proportional to the field

$$\mathbf{p} = \alpha\epsilon_0\mathbf{E} \tag{42.3}$$

The quantity α is called the *molecular electric susceptibility*. Substituting this expression in equation (42.2), we find

$$\mathbf{P} = \alpha\epsilon_0\mathbf{E}\frac{1}{\Delta V}\sum_i 1 = \alpha\epsilon_0\mathbf{E}N \tag{42.4}$$

taking into account that

$$\sum_i 1 = \Delta V N$$

Here N is the number of molecules per unit volume. The molecular electric

susceptibility α is related to the corresponding quantity in the Gaussian system of units α' by

$$\alpha = 4\pi\alpha'$$

Comparing (42.4) and (14.7), we are able to establish a relationship between the molecular electric susceptibility α, and the electric susceptibility κ

$$\kappa = \alpha N \tag{42.5}$$

Hence, from equation (14.25), we obtain the following expression for the permittivity

$$\frac{\epsilon}{\epsilon_0} = \epsilon' = 1 + \alpha N \tag{42.6}$$

It is clear that the molecular electric susceptibility is governed by the internal properties of the molecule and cannot depend much on the density of the substance, nor on pressure and temperature. Hence, for nonpolar molecules, the electric susceptibility of molecules is a linear function of the density of the substance. It can only depend on temperature implicitly, through the value of N. The quantity

$$\frac{\epsilon' - 1}{\rho_m} \tag{42.7}$$

where ρ_m is the density of the substance, is constant and independent of temperature.

Theory of the Polarization of Dielectrics Consisting of Polar Molecules. Many molecules possess permanent electrical moments. For example, the permanent moment of the HCl molecule is approximately 3.44×10^{-30} coul m. That of HBr is 2.33×10^{-30} coul m, etc. Usually the dipole moments of molecules are of the order of 10^{-28} to 10^{-29} coul m. A dipole moment $\mathbf{p}$ in an electric field E has the potential energy

$$W = -\mathbf{p} \cdot \mathbf{E} \tag{42.8}$$

This energy reaches its minimum value when the direction of the dipole coincides with the direction of the electric field. Since the state of minimum energy is the stable state of a system, the dipole moments of polar molecules will tend to turn so as to coincide with the direction of the electric field. This turning is produced by a couple acting on the positive and negative charges of the dipole (Fig. 18). However, thermal motion disrupts this ordering action of the electric field, and, as a result, a different state of equilibrium is established.

Let us take the z axis in the direction of the electric field (Fig. 55). Equation (42.8) states that the energy of the molecules is determined by

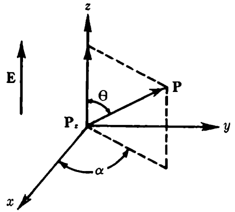


Fig. 55

the angle θ between the z axis and the direction of their dipole moments

$$W = -pE\cos\theta = -p_z E \tag{42.9}$$

Thus, Boltzmann's equation, which describes the energy distribution of particles, becomes, in the present case, an equation describing the distribution of the molecular dipoles with respect to their angles. The number of molecules dN with dipole moments within the solid angle $d\Omega$ is

$$dN = N_0 e^{\frac{pE\cos\theta}{kT}}\, d\Omega = N_0 e^{\frac{pE\cos\theta}{kT}}\, d\alpha \sin\theta\, d\theta \tag{42.10}$$

Hence, the following expression is obtained for the mean value of the projection of the dipole moment on the z axis

$$\langle p_z\rangle = \frac{p\int_0^\pi e^{\beta\cos\theta}\cos\theta\sin\theta\, d\theta}{\int_0^\pi e^{\beta\cos\theta}\sin\theta\, d\theta} \tag{42.11}$$

where

$$\beta = \frac{pE}{kT}$$

The problem reduces to the evaluation of the integral

$$I = \int_0^\pi e^{\beta\cos\theta}\sin\theta\, d\theta \tag{42.12}$$

since the integral in the numerator of (42.11) is easily evaluated by differentiation

$$\int_0^\pi e^{\beta\cos\theta}\cos\theta\sin\theta\, d\theta = \frac{\partial I}{\partial\beta} \tag{42.13}$$

The integral (42.12) is evaluated directly

$$I = \int_0^\pi e^{\beta\cos\theta}\sin\theta\, d\theta = -\frac{1}{\beta}e^{\beta\cos\theta}\Big|_0^\pi = \frac{2}{\beta}\sinh\beta \tag{42.14}$$

Hence, it follows that

$$\frac{\partial I}{\partial \beta} = \frac{2}{\beta}\left(\cosh \beta - \frac{\sinh \beta}{\beta}\right) \tag{42.15}$$

Thus, using (42.14) and (42.15), equation (42.11) becomes

$$\langle p_z \rangle = pL(\beta) \tag{42.16}$$

where $L(\beta) = \coth \beta - 1/\beta$ is known as the *Langevin function*. The form of this function is shown in Fig. 56.

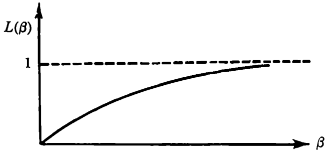


Fig. 56

In practice, we usually meet with fields that are fairly weak, for which $pE \ll kT$, i.e., $\beta \ll 1$. In this case, expanding the hyperbolic cotangent in series

$$\coth \beta = \frac{1}{\beta} + \frac{\beta}{3} - \frac{\beta^3}{45} + \cdots \tag{42.17}$$

we may limit ourselves to the linear term in the expression for $L(\beta)$

$$L(\beta) = \frac{\beta}{3} \tag{42.18}$$

Consequently, we have

$$\langle p_z \rangle = \frac{p^2 E}{3kT} \tag{42.19}$$

Hence, by a process of calculation completely analogous to (42.3) to (42.6), we obtain the following expression for the permittivity

$$\frac{\epsilon}{\epsilon_0} = \epsilon' = 1 + \frac{p^2}{3kT\epsilon_0} N \tag{42.20}$$

Thus, in the case of polar molecules in weak electric fields, the permittivity depends explicitly on temperature.

In very strong fields when $pE \gg kT$, i.e., $\beta \gg 1$, $L(\beta)$ is close to unity

$$L(\beta)\Big|_{\beta \gg 1} \approx 1 \tag{42.21}$$

In such fields

$$\langle p_z \rangle = p \tag{42.22}$$

i.e., all the dipole moments are oriented along the field, and the polarization attains its maximum (saturation) value. Taking the order of magnitude of the dipole moments to be 10^{-29} coul m, we conclude that at $T = 300°$ K, the value of E which produces saturation is equal to

$$E_{\text{sat}} \approx \frac{kT}{p} \approx 4.2 \times 10^{8} \text{ V/m}$$

Case of the Simultaneous Presence of Permanent and Induced Dipole Moments. When polar molecules are placed in an electric field, an additional dipole moment is induced by the mechanism discussed in the case of nonpolar molecules. The polarization of the dielectric due to this induced dipole moment is superimposed on the polarization due to the reorientation of the dipole moments. Assuming that the resultant polarization vector may be expressed in the form of the sum of the polarization vectors arising from the permanent and induced dipole moments, and using equations (42.6) and (42.20), we obtain the following expression for the permittivity

$$\frac{\epsilon}{\epsilon_0} = \epsilon' = 1 + N\left(\alpha + \frac{p^2}{3kT\epsilon_0}\right) \tag{42.23}$$

This is the *Langevin-Debye formula.*

The derivation of equation (42.33) is not very rigorous. However, a more rigorous discussion leads to the same formula. The Langevin-Debye formula agrees well with experiment.

§43. Dense Gases, Liquids, and Solid Dielectrics

When the density of a substance is high, the field acting on a molecule may differ considerably from the external field, and we must take into account the field created by the polarized dielectric itself.

Let us first consider dense gases and slightly polar liquids, i.e., liquids in which the principal component of the polarization is due to the moments induced by the external field. If the constant dipole moment of a molecule of a liquid dielectric is $p_0 \gtrsim 0.16 \times 10^{-29}$ coul m, then the induced polarization plays the dominant role, and the liquid is regarded as slightly polar. If, however, $p_0 \gtrsim 0.33 \times 10^{-29}$ coul m, then it is the orientation polarization which plays the leading role, and such liquids are polar. Generally speaking, the results of this section are not applicable to polar liquids.

Calculation of the Local Field Intensity. Let us consider a molecule in a dielectric, and work out the local field acting upon it. Taking this molecule

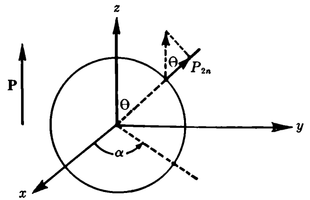


Fig. 57

as the center, we shall draw a physically small sphere around it (Fig. 57). The additional field acting on the molecule consists of two parts: (1) the field $\mathbf{E}_1$ set up by the polarized dielectric outside the sphere; (2) the field $\mathbf{E}_2$ set up by the polarized dielectric inside the sphere.

To calculate $\mathbf{E}_1$, we may use the methods of phenomenological electrodynamics, assuming that the dielectric is a continuous medium. Since the sphere has a physically small volume, we may assume that the medium close to the sphere is uniformly polarized. If we cut out a spherical hollow inside the sphere, then, in accordance with (14.20), we must consider the field $\mathbf{E}_1$ at the center of this hollow to be due to the bound surface charges

$$\sigma_{\text{bound}} = P_{1n} - P_{2n} = -P_{2n} \tag{43.1}$$

Taking the z axis in the direction of the permanent polarization vector $\mathbf{p}$, we obtain

$$\sigma_{\text{bound}} = -P_{2n} = -P\cos\theta \tag{43.2}$$

The surface charge in a solid angle $d\Omega$ is

$$dq = \sigma_{\text{bound}} r^2\, d\Omega \tag{43.3}$$

where r is the radius of the sphere. This charge creates a field in the direction of the z axis,

$$dE_z = -\frac{1}{4\pi\epsilon_0}\frac{dq}{r^2}\cos\theta \tag{43.4}$$

From considerations of symmetry, it is clear that the only nonzero component of the field is the z component. From (43.4) it follows that

$$\begin{aligned} E_z = E_1 &= \frac{1}{4\pi\epsilon_0} P \int \cos^2\theta\, d\Omega \\ &= \frac{1}{4\pi\epsilon_0} P \int_0^{2\pi} d\alpha \int_0^{\pi} \cos^2\theta \sin\theta\, d\theta = \frac{1}{3\epsilon_0} P \end{aligned} \tag{43.5}$$

Consequently, the field set up by the polarized dielectric outside the sphere is equal to

$$\mathbf{E}_1 = \frac{1}{3\epsilon_0}\mathbf{P} \tag{43.6}$$

The value of $\mathbf{E}_2$ is calculated for the special case when the molecules are at the sites of a cubic lattice. We take the origin at the site of the molecule under consideration and the x, y, z axes along the axes of the cubic lattice. To calculate the field we use equation (14.5); the x component of the field is

$$E_{2x} = \frac{p_x}{4\pi\epsilon_0}\sum_i \frac{-r_i^2 + 3x_i^2}{r_i^5} + \frac{p_y}{4\pi\epsilon_0}\sum_i \frac{3x_iy_i}{r_i^5} + \frac{p_z}{4\pi\epsilon_0}\sum_i \frac{3x_iz_i}{r_i^5} \tag{43.7}$$

Here the summation is carried out over all molecules in the physically small volume within the sphere. We may first of all sum over all molecules at a given distance r, and then over the spherical cells corresponding to different distances. In the first summation, in view of the cubic symmetry, we obtain

$$\begin{gathered} \sum_i x_i^2 = \sum_i y_i^2 = \sum_i z_i^2 = \frac{1}{3}\sum_i r_i^2 \\ \sum_i x_iy_i = \sum_i y_iz_i = \sum_i z_ix_i = 0 \end{gathered} \tag{43.8}$$

Hence, from (43.7) and (43.8), it follows that

$$E_{2x} = 0 \tag{43.9}$$

Similarly, we prove that $E_{2y} = E_{2z} = 0$. Hence

$$\mathbf{E}_2 = 0 \tag{43.10}$$

Thus, the local field $\mathbf{E}'$ acting on a molecule within a dielectric is equal to

$$\mathbf{E}' = \mathbf{E} + \frac{1}{3\epsilon_0}\mathbf{P} \tag{43.11}$$

This equation, and the corollaries deduced from it, must be considered as a first approximation since a real dielectric differs from the model used in deducing equation (43.11). In particular, the electric fields of molecules are not simple dipole fields, the lattice of the dielectric may not be cubic, the moments of the molecules may not be identical, etc.

Calculation of the Permittivity. Equations (42.6), (42.20) and (42.23) are obtained from the relationship

$$\epsilon' - 1 = \frac{1}{\epsilon_0}\frac{P}{E} \tag{43.12}$$

where E is the field acting on a molecule. Since the field acting on a mole-

cule in a dense medium is given by equation (43.11), we must replace the quantity (43.12) in these equations by

$$\frac{1}{\epsilon_0}\frac{P}{E + \dfrac{1}{3\epsilon_0}P} = \frac{\kappa}{1 + \dfrac{1}{3}\kappa} = \frac{3(\epsilon' - 1)}{\epsilon' + 2} \tag{43.13}$$

where we use $P = \kappa\epsilon_0 E$, $\kappa = \epsilon' - 1$. Hence, the Langevin-Debye formula (42.23) in the approximation under consideration takes the form

$$\frac{3(\epsilon' - 1)}{\epsilon' + 2} = N\left(\alpha + \frac{p^2}{3kT\epsilon_0}\right) \tag{43.14}$$

This is often called the *Clausius-Mossotti formula.*

Equation (43.14) is valid for gases of low density and for nonpolar liquids. For polar liquids and solids it gives incorrect results. Moreover, the temperature dependence of ϵ' given by this equation is not confirmed by experiment even for slightly polar liquids. In spite of this, equation (43.14) may be considered as the first rough approximation. For rarefied gases, in which $\epsilon' \approx 1$, this equation, as would be expected, transforms into equation (42.23).

If we ignore the polar nature of a dielectric, $p \approx 0$, then equation (43.14) may be written in the form

$$\frac{3(\epsilon' - 1)}{\epsilon' + 2} = \alpha N = \frac{\rho_m\alpha}{M} \tag{43.15}$$

where ρ_m is the density of the dielectric, and M is the mass of a molecule of the dielectric. Thus, for nonpolar dielectrics the value of

$$\frac{\epsilon' - 1}{(\epsilon' + 2)\rho_m} = \frac{\alpha}{3M} = \text{const} \tag{43.16}$$

is independent of temperature and density. For polar dielectrics, the value of

$$\frac{\epsilon' - 1}{(\epsilon' + 2)\rho_m} = \frac{1}{3M}\left(\alpha + \frac{p^2}{3kT\epsilon_0}\right) \tag{43.17}$$

depends on temperature.

Solid Dielectrics and Polar Liquids. Comparison of equation (43.14) with experiment shows that it describes gases and nonpolar liquids in a satisfactory manner. However, in the case of polar liquids, this equation often leads to absurd results. For example, the dipole moment of the water molecule is 0.62×10^{-29} coul m, and the density at 20° C is 0.99 g/cm^3. Hence, from (43.14), the relative permittivity of water has a negative value $\epsilon' = -2.83$, while the experimental value is $\epsilon' = 81$. In general, equation (43.14) gives negative permittivity for polar liquids with per-

manent molecular dipole moments $p \gtrsim 0.5 \times 10^{-29}$ coul m. This equation gives results which agree satisfactorily with experiment only for liquids with small dipole moment. The reason why this equation is not applicable to polar liquids is the fact that it does not take into account the strong interactions between molecules. The equations obtained above also fail to give good results when applied to solid dielectrics. The deviation from these equations in the case of solid dielectrics is due to the following causes.

Nonpolar molecular crystals, which have nonpolar molecules at crystal lattice sites, behave like nonpolar liquids, i.e., their permittivity is independent of temperature. However, the quantitative equations obtained in the present section for nonpolar liquids do not give correct results for nonpolar molecular crystals, since in a crystal the molecules are fixed at their sites in the crystal lattice and can only oscillate about their equilibrium positions, while in a liquid they are able to move about more freely. Hence, static averaging over the positions of the molecules in a crystal is somewhat different from that in the case of a liquid. When the thermal oscillations of the molecules are negligibly small, and the relative permittivity is close to unity, then the equations obtained above for nonpolar liquids are also a good first approximation in the case of nonpolar crystals.

In polar crystals, which have polar molecules at the sites of the crystal lattice, the intermolecular forces have a considerable effect on the orientation of the molecules. Hence, equation (43.14) cannot be applied to polar crystals, especially at low temperatures.

In ionic crystals, which have ions at the sites of the crystal lattice, it is often not the displacement of the electrons within atoms that plays the principal part in the polarization, but the relative displacement of the positive and negative ions, which must be considered separately.

Thus, the theory discussed above requires considerable refinement in the case of polar liquids and solids, and such a refinement is outside the scope of this book.

§44. Theory of Dispersion

As equation (34.21) states, the refractive index of a dielectric with respect to empty space is given by

$$n = \sqrt{\frac{\epsilon}{\epsilon_0}} = \sqrt{\epsilon'} \tag{44.1}$$

where ϵ_0 is the permittivity of empty space. The expressions for the permittivity obtained in the previous paragraph give satisfactory agreement with the experimental data in the case of steady or slowly changing electromagnetic fields. However, in the case of rapidly changing electromagnetic

fields, in particular, for electromagnetic waves in the region of visible light, the refractive index depends on the frequency. The dependence of the refractive index on the frequency is called the *dispersion*. The dispersion of light, for example, is demonstrated by the resolution of white light into the colors of the spectrum when it passes through a prism.

Forced Oscillations of an Elastically Bound Electron. When an electron in an atom moves away from its equilibrium position, a restoring force tends to make it return to the equilibrium position. If the displacement from the equilibrium position is small, then the restoring force may be expanded as a series in terms of the displacement, and we may take only the first term of the expansion, which represents an elastic force proportional to the displacement (*Hooke's law*). Under the action of such a force, the electron performs elastic oscillations of frequency ω_0. The equation for free oscillations of the electron, with allowance for damping, is of the form

$$m\ddot{\mathbf{r}} + m\gamma\dot{\mathbf{r}} + m\omega_0^2\mathbf{r} = 0 \tag{44.2}$$

where the quantity γ represents the damping.

If a monochromatic electromagnetic wave falls on an atom, then, under the action of the electric vector of the wave, the electron performs forced oscillations given by

$$m\ddot{\mathbf{r}} + m\gamma\dot{\mathbf{r}} + m\omega_0^2\mathbf{r} = e\mathbf{E}_0e^{i\omega t} \tag{44.3}$$

where E_0 and ω are the amplitude and frequency of the incident wave. The solution of (44.3) is written in the form

$$\mathbf{r} = \frac{e}{m}\frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega}\mathbf{E} \qquad \mathbf{E} = \mathbf{E}_0e^{i\omega t} \tag{44.4}$$

Polarization of a Dielectric. Displacement of a charge e to a distance $\mathbf{r}$ produces a dipole moment $\mathbf{p}$

$$\mathbf{p} = e\mathbf{r} = \alpha\epsilon_0\mathbf{E}$$

$$\alpha = \frac{e^2}{\epsilon_0 m}\frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega} \tag{44.5}$$

Repeating the discussion for rarefied gases in connection with the derivation of equations (42.3) to (42.6), we obtain the following expression for the permittivity in the present case

$$\frac{\epsilon_\omega}{\epsilon_0} = \epsilon'_\omega = 1 + \alpha N = 1 + \frac{e^2}{\epsilon_0 m}\frac{N}{\omega_0^2 - \omega^2 + i\gamma\omega} \tag{44.6}$$

In the case of dense gases and condensed media, a local field, defined by (43.11), acts on the electron. Repeating the discussion which has led to equation (43.14), we obtain

$$\frac{3(\epsilon'_\omega - 1)}{\epsilon'_\omega + 2} = \alpha N \tag{44.7}$$

where α is defined in (44.5). Solving this equation for ϵ'_ω, we find

$$\frac{\epsilon_\omega}{\epsilon_0} = \epsilon'_\omega = 1 + \frac{e^2}{\epsilon_0 m} \frac{N}{\omega_0^2 - \omega^2 + i\gamma\omega - \frac{1}{3}\frac{e^2}{\epsilon_0 m} N} \tag{44.8}$$

The expression obtained has the form of (44.6), with the square of the frequency ω_0^2 replaced by

$$\omega_0^2 - \frac{1}{3}\frac{e^2}{\epsilon_0 m} N$$

Otherwise, equations (44.6) and (44.8) are identical.

Equation (44.6) states that ϵ'_ω is close to unity when

$$\frac{e^2}{\epsilon_0 m} N \ll |\omega_0^2 - \omega^2| \tag{44.9}$$

This applies to rarefied gases. But, in this case, as one would expect, equation (44.8) is transformed into equation (44.6).

Equations (44.6) and (44.8) show that the permittivity is a complex quantity. Consequently, the refractive index n'_ω, related to the relative permittivity ϵ'_ω by (44.1), is also a complex quantity. Using equation (44.6) for ϵ'_ω, we obtain

$$n'^2_\omega = \epsilon'_\omega = 1 + \frac{e^2}{\epsilon_0 m} \frac{N}{\omega_0^2 - \omega^2 + i\gamma\omega} \tag{44.10}$$

Hence, it follows that if n'_ω has the form

$$n'_\omega = n_\omega - i\xi_\omega \tag{44.11}$$

then

$$n_\omega^2 - \xi_\omega^2 = 1 + \frac{e^2 N}{\epsilon_0 m} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{44.12a}$$

$$2n_\omega\xi_\omega = \frac{e^2 N}{\epsilon_0 m} \frac{\omega\gamma}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{44.12b}$$

The presence of the imaginary part in the complex refractive index is associated with the absorption of the electromagnetic wave. The real part describes the phenomenon of dispersion.

Normal Dispersion in the Optical Region. In the optical region

$$\gamma^2\omega^2 \ll (\omega_0^2 - \omega^2)^2 \tag{44.13}$$

and, hence, we may take $\xi_\omega \approx 0$. Equation (44.12a) in this case takes the form

$$n_\omega^2 = 1 + \frac{e^2}{\epsilon_0 m} \frac{N}{\omega_0^2 - \omega^2} \tag{44.14}$$

This equation is obtained on the assumption that there are N electrons per unit volume of natural frequency of oscillation ω_0. In fact, however, not all electrons in an atom are held in the position of equilibrium by the same elastic forces, and, hence, they do not all have the same natural frequency. Let us denote by N_i the number of electrons per unit volume which have the natural frequency ω_{0i}. In this case, equation (44.14) may be generalized

$$n_\omega^2 = 1 + \frac{e^2}{\epsilon_0 m} \sum_i \frac{N_i}{\omega_{0i}^2 - \omega^2} \tag{44.15}$$

If the refractive index is close to unity (which is true for sufficiently rarefied gases, i.e., when $n_\omega - 1 \ll 1$, it follows from (44.15) that

$$n_\omega = 1 + \frac{e^2}{2\epsilon_0 m} \sum_i \frac{N_i}{\omega_{0i}^2 - \omega^2} \tag{44.16}$$

Fig. 58 gives the general form of the dispersion curve. Throughout the whole transparent region, the refractive index increases as the frequency

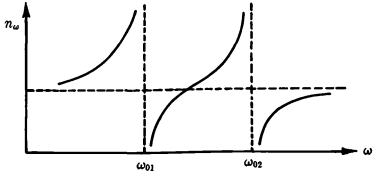


Fig. 58

increases. At low frequencies $\omega \ll \omega_{0i}$, equation (44.16) gives the static value of the refractive index

$$n = 1 + \frac{e^2}{2\epsilon_0 m} \sum_i \frac{N_i}{\omega_{0i}^2} \tag{44.17}$$

At very high frequencies $\omega \gg \omega_{0i}$, the refractive index tends to unity but remains less than unity, since all terms in the sum (44.16) are negative. This means that for ultrashort waves, the dielectric is a less optically dense medium than empty space, so that total reflection may occur. This is observed in the case of x-rays.

In the case of hard x-rays, $\omega \gg \omega_{0i}$, equation (44.6) takes the form

$$n_\omega = 1 - \frac{e^2}{2\epsilon_0 m} \frac{1}{\omega^2} \sum_i N_i \tag{44.18}$$

Thus, in this case the nature of the binding between the electrons and atoms plays no part at all, and refractive index is determined only by the total number of electrons.

Anomalous Dispersion. The dispersion curve in Fig. 58 is obtained from (44.15), which has been deduced from (44.6), ignoring the damping ($\gamma = 0$). If the damping is taken into account, then at the point $\omega = \omega_{0i}$, the dispersion curve changes continuously, and not as shown in Fig. 58. Let us represent the refractive index in the form (44.11). If $|n'_\omega|$ does not differ greatly from unity, then using (44.6), we obtain

$$n_\omega - i\xi_\omega = \sqrt{\epsilon'_\omega} = 1 + \frac{e^2}{2\epsilon_0 m} \frac{N}{\omega_0^2 - \omega^2 + i\gamma\omega} \tag{44.19}$$

Hence, it follows that

$$n_\omega = 1 + \frac{e^2}{2\epsilon_0 m} N \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{44.20a}$$

$$\xi_\omega = \frac{e^2}{2\epsilon_0 m} N \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{44.20b}$$

The behavior of the dispersion curve close to the resonance frequency is shown in Fig. 59. Close to the resonance frequency ω_0, the refractive index

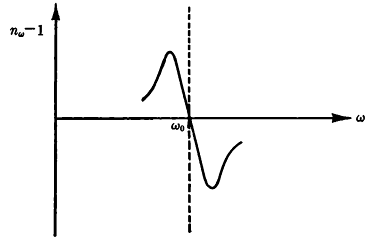


Fig. 59

decreases with increase of the frequency. This decrease is called *anomalous dispersion.*

Absorption. In §32, it has been shown that if the electric vector of a wave traveling in the z direction is of the form (32.12)

$$\mathbf{E}(z, t) = \mathbf{E}_0 e^{i(\omega t - kz)} \tag{44.21}$$

then the wave number k is related to the permittivity and permeability by (32.8)

$$k = \omega\sqrt{\epsilon\mu} = \frac{\omega}{c}\sqrt{\epsilon'\mu'} \tag{44.22}$$

Since for dielectrics $\mu' \approx 1$, substituting in (44.22) the expression for $\sqrt{\epsilon'_\omega}$ from (44.19), we obtain

$$k = \frac{\omega}{c}n_\omega - i\frac{\omega}{c}\xi_\omega \tag{44.23}$$

Using this expression for k, equation (44.21) becomes

$$\mathbf{E}(z, t) = \mathbf{E}_0 e^{-\frac{\omega}{c}\xi_\omega z} e^{i\left(\omega t - \frac{\omega}{c}n_\omega z\right)} \tag{44.24}$$

Thus, the imaginary part of the refractive index describes the damping of a plane wave in a dielectric. This damping is caused by the fact that when a wave passes through a dielectric, it does work on the individual atoms of the dielectric: the energy of the wave is used up in initiating forced oscillations of the electrons. The oscillating electrons dissipate this energy in all directions as radiation energy.

PROBLEMS

1 Calculate the relative permittivity of helium under normal conditions ($t = 15°$ C, $p = 1$ atm) if the atomic polarizability for helium equals

$$\alpha = 2.48 \times 10^{-30}\ \mathrm{m}^3$$

Solution: Loschmidt's number is $L = 2.7 \times 10^{25}\ \mathrm{m}^{-3}$. Consequently

$$\epsilon' = 1 + L\alpha = 1 + 6.7 \times 10^{-5} = 1.000067$$

(the experimental value $\epsilon' = 1.000074$)

2 Calculate the permittivity of ammonia at $t = 27°$ C, $\alpha = 1.37 \times 10^{-29}\ \mathrm{m}^3$, $p = 0.46 \times 10^{-29}$ coul m.
Hint: Use equation (42.23).

Answer: $\epsilon = 1.0076\epsilon_0$

3 Find the electric susceptibility α for the hydrogen atom. The electric field is taken perpendicular to the plane of motion of the electron.
Solution: We shall write down the equilibrium condition for a moving electron in an external field (Fig. 60)

$$eE = \frac{e^2}{4\pi\epsilon_0(x^2 + r^2)}\cos\alpha = \frac{e^2}{4\pi\epsilon_0}\frac{x}{(x^2 + r^2)^{3/2}}$$

i.e.

$$ex = p = 4\pi\epsilon_0 r^3 E$$

since $x \ll r$, and, hence, $(x^2 + r^2)^{3/2} \approx r^3$, it follows that

$$\alpha = 4\pi r^3 \approx 1.57 \times 10^{-30}\ \mathrm{m}^3$$

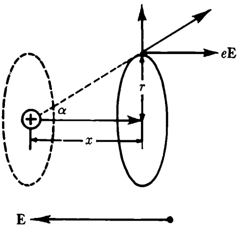


Fig. 60

This equation gives the correct order of the value of the atomic electric susceptibility.

CHAPTER 9

Magnetic Substances

§45. Motion of Electrons in Atoms in an External Magnetic Field

Substances which become magnetized in the presence of a magnetic field are called magnetic substances. In order to consider the various methods of magnetization, we must first consider the motion of electrons in atoms in a magnetic field. Let us take the origin of coordinates at the center of the nucleus of an atom and the z axis in the direction of the magnetic field. In this coordinate system

$$B_x = B_y = 0 \qquad B_z = B \tag{45.1}$$

When there is no external magnetic field, every electron in the atom moves in the electric field of the nucleus and the other electrons. Since the electron velocities in the atom are nonrelativistic, the magnetic interaction between the electrons may be ignored. We shall denote the coordinates of the k^{th} electron by x_k, y_k, z_k. The potential function takes the form

$$\Phi(x_1, \ldots, x_n, y_1, \ldots, y_n, z_1, \ldots, z_n) = \sum \varphi(r_i) + \sum \frac{e}{r_{ik}} \tag{45.2}$$

where $r_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$ is the distance of the i^{th} electron from the nucleus, $r_{ik} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2 + (z_i - z_k)^2}$ is the distance between the i^{th} and k^{th} electrons.

The first term on the right-hand side of (45.2) represents the interaction of the electrons with the nucleus, while the second term represents the mutual interaction of the electrons. The equations of motion of the k^{th} electron in the absence of an external magnetic field take the form

$$m\ddot{x}_k = -e\frac{\partial \Phi}{\partial x_k} \qquad m\ddot{y}_k = -e\frac{\partial \Phi}{\partial y_k} \qquad m\ddot{z}_k = -e\frac{\partial \Phi}{\partial z_k} \tag{45.3}$$

The solutions of these equations are given by some functions

$$x_k = x_k(t) \qquad y_k = y_k(t) \qquad z_k = z_k(t) \tag{45.4}$$

In the presence of an external magnetic field, an additional Lorentz force acts on each electron

$$\mathbf{F} = e\mathbf{v} \times \mathbf{B} \tag{45.5}$$

The components of the Lorentz force acting on the k^{th} electron are equal to

$$\begin{aligned} F_{kx} &= e(v_{ky}B_z - v_{kz}B_y) = e\dot{y}_k B \\ F_{ky} &= e(v_{kz}B_x - v_{kx}B_z) = -e\dot{x}_k B \\ F_{kz} &= e(v_{kx}B_y - v_{ky}B_x) = 0 \end{aligned} \tag{45.6}$$

Using primes to denote the coordinates of an electron in the presence of a magnetic field, and taking (45.6) into consideration, we may write down the following equations of motion

$$\begin{aligned} m\ddot{x}'_k &= -e\frac{\partial \Phi'}{\partial x'_k} + eB\dot{y}'_k \\ m\ddot{y}'_k &= -e\frac{\partial \Phi'}{\partial y'_k} - eB\dot{x}'_k \\ m\ddot{z}'_k &= -e\frac{\partial \Phi'}{\partial z'_k} \end{aligned} \tag{45.7}$$

where $\Phi' = \Phi(x'_1, \ldots, x'_n, y'_1, \ldots, y'_n, z'_1, \ldots, z'_n)$. Since the Lorentz force (45.5) has no component in the direction of the magnetic field, the latter has no effect on the component of velocity along the magnetic field. Hence, we have

$$z'_k = z_k(t) \tag{45.8}$$

It is convenient to study the motion in the xy plane by introducing the complex variable

$$\zeta'_k = x'_k + iy'_k \tag{45.9}$$

Multiplying the second equation of (45.7) by i, and adding it to the first equation, we obtain the following equation for ζ'_k

$$m\ddot{\zeta}'_k = -e\left(\frac{\partial \Phi'}{\partial x'_k} + i\frac{\partial \Phi'}{\partial y'_k}\right) - ieB\dot{\zeta}'_k \tag{45.10}$$

In the absence of a magnetic field, the equation for

$$\zeta_k = x_k + iy_k \tag{45.11}$$

takes the form

$$m\ddot{\zeta}_k = -e\left(\frac{\partial \Phi}{\partial x_k} + i\frac{\partial \Phi}{\partial y_k}\right) \tag{45.12}$$

Let us seek the solution of (45.10) in the form

$$\zeta'_k = \xi_k e^{i\omega t} \tag{45.13}$$

Substituting (45.13) into (45.10), we obtain the following equation for ξ_k

$$m\ddot{\xi}_k = -e\left(\frac{\partial\Phi'}{\partial x'_k} + i\frac{\partial\Phi'}{\partial y'_k}\right)e^{-i\omega t} - i\dot{\xi}_k(eB + 2m\omega) + \xi_k(eB\omega + m\omega^2) \tag{45.14}$$

We choose the frequency ω in such a way that the term with $\dot{\xi}_k$ becomes zero

$$eB + 2m\omega = 0 \qquad \omega = \omega_L = -\frac{eB}{2m} \tag{45.15}$$

Then equation (45.14) becomes

$$m\ddot{\xi}_k = -e\left(\frac{\partial\Phi'}{\partial x'_k} + i\frac{\partial\Phi'}{\partial y'_k}\right)e^{i\omega_L t} - m\omega_L{}^2\xi_k \tag{45.16}$$

We shall assume that the frequency ω_L is such that the last term in (45.16) may be ignored. It is asserted that in this case the solution of the equation

$$m\ddot{\xi}_k = -e\left(\frac{\partial\Phi'}{\partial x'_k} + i\frac{\partial\Phi'}{\partial y'_k}\right)e^{-i\omega_L t} \tag{45.17}$$

is identical with the solution of (45.12) in the absence of a magnetic field, i.e.

$$\xi_k = \zeta_k \tag{45.18}$$

To verify this, we note that the transformation (45.13)

$$\zeta'_k = \zeta_k e^{i\omega_L t} \tag{45.19}$$

describes the rotation of the vector ζ_k through an angle $\omega_L t$. This means that the electron coordinates with a prime are obtained from the coordinates without a prime by rotation of all electrons in the atom through an angle $\omega_L t$ about the z axis. When such a rotation takes place, the relative distances of the electrons r_{ik} in the expression for Φ are invariant, as are also the values of r_i. Hence, the complex force

$$-e\left(\frac{\partial\Phi}{\partial x_k} + i\frac{\partial\Phi}{\partial y_k}\right) \tag{45.20}$$

also rotates, as a result of the transformation (45.19), through an angle $\omega_L t$ without changing in absolute magnitude. Hence, the complex force

$$-e\left(\frac{\partial\Phi'}{\partial x'_k} + i\frac{\partial\Phi'}{\partial y'_k}\right) \tag{45.21}$$

is obtained from the complex force (45.20) by turning through an angle $\omega_L t$, i.e.

$$-e\left(\frac{\partial\Phi'}{\partial x'_k} + i\frac{\partial\Phi'}{\partial y'_k}\right) = -e\left(\frac{\partial\Phi}{\partial x_k} + i\frac{\partial\Phi}{\partial y_k}\right)e^{i\omega_L t} \tag{45.22}$$

Taking (45.22) into account, we may rewrite (45.17) in the form

$$m\ddot{\xi}_k = -e\left(\frac{\partial\Phi}{\partial x_k} + i\,\frac{\partial\Phi}{\partial y_k}\right) \tag{45.23}$$

which agrees with (45.12). This proves the validity of (45.18).

Thus, when an atom is in a magnetic field, all its electrons receive an additional angular velocity about the direction of the magnetic field with the frequency

$$\omega_L = \frac{|e|B}{2m} \tag{45.24}$$

which is called the *Larmor frequency*. This is the essence of Larmor's theorem.

We must now find the magnetic fields in which this proof of the theorem holds. It holds when the last term on the right-hand side of (45.16) may be ignored in comparison with the other terms. The magnitude of the main term in this equation is of the order of

$$m\ddot{\xi}_k \approx m\omega_0{}^2\xi_k \tag{45.25}$$

where ω_0 is the frequency of the periodic motion of the electron in the atom, and has the same order of magnitude as the frequencies of visible light, i.e., $\omega_0 \approx 10^{15}\ \mathrm{sec}^{-1}$. Hence, the condition of validity of the proof may be written

$$\frac{\omega_L}{\omega_0} = \frac{|e|B}{2m\omega_0} \ll 1 \tag{45.26}$$

This is satisfied when

$$B \ll \frac{2m\omega_0}{|e|} \approx 10^4\ \mathrm{T} = 10^8\ \mathrm{G} \tag{45.27}$$

A field of 10^8 G is considerably stronger than the magnetic fields used in practical work. Hence, (45.27) is satisfied in all cases of practical interest.

Precession of Atoms. If an electron moves on a closed orbit in the central field of the nucleus, then the magnetic moment of the closed current created by the electron is equal to

$$M = \frac{|e|}{T}S \tag{45.28}$$

where S is the area enclosed by the orbit and T is the period of revolution.

The angular momentum L is conserved in the case of motion in a central force field

$$L = mr^2\frac{d\varphi}{dt} = \mathrm{const} \tag{45.29}$$

where r, φ are the polar coordinates of the electron in the plane of motion.

The area enclosed by the orbit of the moving electron is, in polar coordinates

$$S = \frac{1}{2}\int_0^{2\pi} r^2\, d\varphi \tag{45.30}$$

Transforming (45.30), integrating with respect to time, and taking (45.29) into account, we obtain

$$S = \frac{1}{2}\int_0^T r^2 \frac{L}{mr^2}\, dt = \frac{1}{2}\frac{L}{m} T \tag{45.31}$$

Hence, equation (45.28) becomes

$$M = \frac{|e|}{2m} L \tag{45.32}$$

The magnetic moment and the angular momentum are vectors. For a positively charged particle, they act in the same direction, and for a negatively charged particle, they act in opposite directions. Hence, if e represents the charge of a particle, taking its sign into account, we can write down the following expression for the relationship between the angular momentum and the magnetic moment of a particle moving around an orbit in a central force field

$$\mathbf{M} = \frac{e}{2m}\mathbf{L} \tag{45.32a}$$

As was shown in §22, a magnetic moment in a magnetic field experiences a couple $\mathbf{N}$, which tends to align the direction of the magnetic moment along the direction of the magnetic field (Fig. 61)

$$\mathbf{N} = \mathbf{M} \times \mathbf{B} \tag{45.33}$$

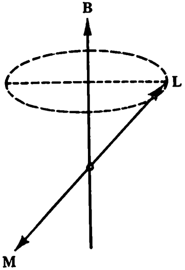


Fig. 61

In addition to the magnetic moment, the atom also possesses angular momentum. Under the action of the couple N of (45.33), which tends to change the angle between **M** and **B**, the angular momentum begins to precess about the direction of the magnetic field, as in any gyroscope on attempting to alter the direction of the momentum. The equation of motion of a gyroscope under the action of a couple is well known from theoretical mechanics, and takes the form

$$\frac{d\mathbf{L}}{dt} = \mathbf{N} = \mathbf{M} \times \mathbf{B} \tag{45.34}$$

Substituting the expression for **M** from (45.32′), we may rewrite (45.34) in the form

$$\frac{d\mathbf{L}}{dt} = \frac{e}{2m}\mathbf{L} \times \mathbf{B} = -\frac{e}{2m}\mathbf{B} \times \mathbf{L} = \boldsymbol{\omega}_L \times \mathbf{L} \tag{45.35}$$

where

$$\boldsymbol{\omega}_L = -\frac{e}{2m}\mathbf{B}$$

Let us now compare equation (45.35) with the equation of motion of an absolutely rigid body, rotating with an angular velocity $\boldsymbol{\omega}$

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r} \tag{45.36}$$

The comparison shows that equation (45.35) describes the precession of the vector **L** about the direction of **B** with an angular velocity

$$\boldsymbol{\omega}_L = -\frac{e}{2m}\mathbf{B} \tag{45.37}$$

which is called the Larmor frequency.

Thus, the angular momenta of atoms in a magnetic field precess about the direction of the magnetic field at the Larmor frequency. This precessional motion is superimposed on the motion of the atomic electrons and alters the frequency of the electron motion in a plane perpendicular to the direction of the magnetic field. All atoms precess in the same direction. The frequency of the precession is either added to or subtracted from the frequency of revolution of the electrons, depending on whether the direction of revolution of the electrons in the atom is the same as, or opposite to, the direction of precession.

Law of Conservation of Energy. Since the angular velocity of electrons in an atom in a magnetic field varies, their energy will also vary. However, the magnetic field itself performs no work and does not change the energy of charged particles moving in it. What then alters the energy of the electrons? The energy changes due to the work done by the electric field

induced by the change in the magnetic field. This may be confirmed by a calculation.

When an atom is placed in a magnetic field, the change in the energy of the electron is equal to

$$\Delta W = \frac{m}{2} r^2(\omega_0 + \omega_L)^2 - \frac{m}{2} r^2\omega_0^2 \approx mr^2\omega_0\omega_L \tag{45.38}$$

taking into account that $\omega_L \ll \omega_0$, and discarding terms of second-order in ω_L. On the other hand, by Faraday's law of electromagnetic induction, the work done by the induced electric field when a charge e moves round a closed contour L is equal to

$$\Delta W_{\text{rev}} = \left| e \int_L \mathbf{E}\cdot d\mathbf{l} \right| = |e| \left| \frac{d\Phi}{dt} \right| \tag{45.39}$$

Since variations in the magnetic field at distances of the order of atomic dimensions during an interval of time T (the period of revolution of the electron) may be ignored, we may assume that

$$\Delta W_{\text{rev}} = |e|\pi r^2 \frac{dB}{dt} \tag{45.40}$$

Hence, the rate of change of the energy of the electron equals

$$\frac{\Delta W_{\text{rev}}}{T} = \frac{dW}{dt} = \frac{1}{2} |e| r^2\omega_0 \frac{dB}{dt} \tag{45.41}$$

It follows that the total change of the electron energy in the time interval from 0 to t, during which the magnetic field changes from 0 to B, is equal to

$$\Delta W = \int_0^t \frac{dW}{dt}\, dt = \frac{1}{2} |e| r^2\omega_0 \int_0^t \frac{dB}{dt}\, dt = \frac{1}{2} |e| r^2\omega_0 B = mr^2\omega_0\omega_L \tag{45.42}$$

which is identical with equation (45.38). It is thus proved that the change in the energy of motion of electrons when an atom is placed in a magnetic field is, in fact, due to the work done by the induced electric field.

§46. Diamagnetic Substances

Molecular Picture of Diamagnetism. In a magnetic field, all electrons acquire an additional rotation in one and the same direction, and, hence, all atoms acquire an additional magnetic moment, which is opposite in direction to the magnetic field (Fig. 62). As a result of these additional moments, the body becomes magnetized. This method of magnetization

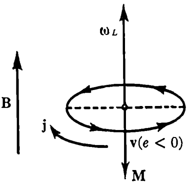


Fig. 62

is known as *diamagnetism*. The magnetization vector **I** is defined as the magnetic moment per unit volume of a dielectric

$$\mathbf{I} = \frac{1}{\Delta V} \sum_i \mathbf{M}_i \tag{46.1}$$

where ΔV is a physically small volume, and the summation is carried out over the magnetic moments $\mathbf{M}_i$ of the atoms in this volume.

Diamagnetic substances are characterized by the fact that in the absence of a magnetic field, the magnetic moments of their atoms are equal to zero, i.e., the atoms of diamagnetic substances do not possess permanent magnetic moments. An external magnetic field produces a change in the angular velocity of the electrons in the atoms, and, hence, induced magnetic moments which cause the diamagnetic material to become magnetized.

Calculation of the Diamagnetic Susceptibility. The *diamagnetic susceptibility* χ_d is the coefficient of proportionality between the magnetization vector **I** and the magnetic field

$$\mathbf{I} = \chi_d \mathbf{H} \tag{46.2}$$

Let us evaluate the magnetic moment of an atom of a diamagnetic substance. We shall take the origin at the center of the atom and the z axis in the direction of the magnetic field (Fig. 63). The velocity $\mathbf{v}_i$ of an electron in the atom is made up of two components, the velocity $\mathbf{v}_{0i}$ which the electron possesses in the absence of a magnetic field, and the velocity $\boldsymbol{\omega}_L \times \mathbf{R}_i$ due to the magnetic field

$$\mathbf{v}_i = \mathbf{v}_{0i} + \boldsymbol{\omega}_L \times \mathbf{R}_i \tag{46.3}$$

The magnetic moment of the atom caused by the motion of the electrons in the absence of a magnetic field is zero for diamagnetic substances. Hence, the first term of (46.3), summed over all electrons, makes no

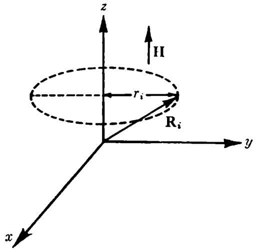


Fig. 63

contribution to the magnetic moment. The magnetic moment is solely due to the second term, and equals

$$M_{zi} = -\frac{1}{2}|e|\omega_L r_i^2 \tag{46.4}$$

where $r_i^2 = x_i^2 + y_i^2$, and the minus sign takes into account the fact that the diamagnetic moment acts in the opposite direction to the magnetic field. To obtain the magnetic moment of the atom, (46.4) must be summed over all electrons of the atom

$$M_z = \sum_i M_{zi} = -\frac{1}{2}|e|\omega_L \sum_i r_i^2 \tag{46.5}$$

We shall now take into account the fact that individual atoms have different arbitrary orientations in space. Equation (46.5) must, therefore, be averaged over these arbitrary orientations

$$\langle M_z \rangle = -\frac{1}{2}|e|\omega_L \langle \sum_i r_i^2 \rangle \tag{46.6}$$

Since

$$\begin{aligned} \langle x_i^2 \rangle = \langle y_i^2 \rangle = \langle z_i^2 \rangle &= \frac{1}{3}\langle R_i^2 \rangle \\ \langle r_i^2 \rangle = \langle x_i^2 \rangle + \langle y_i^2 \rangle &= \frac{2}{3}\langle R_i^2 \rangle \end{aligned} \tag{46.7}$$

we obtain

$$\langle \sum_i r_i^2 \rangle = \sum_i \langle r_i^2 \rangle = \frac{2}{3}\sum \langle R_i^2 \rangle = \frac{2}{3}Z\langle R^2 \rangle \tag{46.8}$$

where $\langle R^2 \rangle$ is the mean distance of the electrons from the nucleus, and Z is the number of electrons in the atom.

Thus, equation (46.6) becomes

$$\langle M_z \rangle = -\frac{1}{3}|e|\omega_L Z \langle R^2 \rangle = -\frac{e^2 Z\mu}{6m}\langle R^2 \rangle H \tag{46.9}$$

Hence, from (46.1), we find

$$I_z = \frac{1}{\Delta V}\sum_j \langle M_{zj} \rangle = \frac{1}{\Delta V}\langle M_z \rangle N\,\Delta V = \langle M_z \rangle N \tag{46.10}$$

where N is the number of atoms per unit volume. Substituting (46.9) in (46.10), we find

$$\mathbf{I} = -\frac{e^2 Z\mu}{6m}\langle R^2 \rangle N\mathbf{H} \tag{46.11}$$

where we return to the vector notation.

Comparison of (46.11) and (46.2) shows that the diamagnetic susceptibility χ_d is equal to

$$\frac{\chi_d}{1+\chi_d} \approx \chi_d \approx -\frac{e^2 Z\langle R^2 \rangle}{6m}\mu_0 N \tag{46.12}$$

taking into account that $\mu = \mu_0(1 + \chi_d)$. However, in diamagnetic substances, $\chi_d \ll 1$ and, consequently, the left-hand side of (46.12) may be taken as equal to χ_d. The minus sign takes into account the fact that in diamagnetic substances, the magnetization vector and the magnetic field act in opposite directions.

Equation (46.12) shows that the electric susceptibility does not explicitly depend on temperature, but depends only on the density of the substance (through the coefficient N, the number of atoms per unit volume).

There is good experimental support of equation (46.12), provided that $\langle R^2 \rangle$ is evaluated according to the methods of the quantum theory of atoms.

§47. Paramagnetic Substances

Molecular Picture of Magnetization. If the magnetic moments of atoms are not equal to zero in the absence of an external magnetic field, reorientation of the atoms occurs when they are placed in a magnetic field. This is because the energy of a magnetic moment $\mathbf{M}$, in a magnetic field $\mathbf{B}$, is given by

$$W = -\mathbf{M}\cdot\mathbf{B} \tag{47.1}$$

Hence, the state of minimum energy will be attained when the magnetic field and the magnetic moment are oriented along the same direction. This orientation of the magnetic moments along the direction of the

magnetic field leads to the magnetization of the substance. This method of magnetization is called *paramagnetism.* In this case, the magnetization vector acts in the direction of the magnetic field, hence, the paramagnetic susceptibility χ_p is positive.

We must note that in paramagnetic substances, there is also a diamagnetic magnetization effect, but the latter is always considerably weaker than the paramagnetic effect. Hence, substances with atoms possessing permanent magnetic moments are *paramagnetic substances.*

Calculation of Paramagnetic Susceptibility. The picture of the magnetization of paramagnetic substances is the same as the picture of the polarization of dielectrics with polar molecules, which has been discussed in detail in §42. The expression for the energy of the magnetic moment (47.1) is analogous in form to the expression for the energy of the dipole moment (42.8). Hence, the theory of the magnetization of paramagnetic substances is completely analogous to the theory of the polarization of dielectrics consisting of polar molecules, when the substitutions $\mathbf{p} \to \mathbf{M}$, $\mathbf{E} \to \mathbf{B}$ are made throughout. Hence, starting with equation (42.8), and repeating the whole discussion word-by-word, we obtain, instead of (42.16)

$$\langle M_{az} \rangle = M_a L(\beta) \qquad \beta = \frac{M_a B}{kT} \tag{47.2}$$

In weak fields, when $M_a B \ll kT$, we have, instead of (42.19)

$$\langle M_{az} \rangle = \frac{M^2}{3kT} B \tag{47.3}$$

Hence, the following expression is obtained for the paramagnetic susceptibility χ_p

$$\frac{\chi_p}{1 + \chi_p} \approx \chi_p = \frac{M_a{}^2 N \mu_0}{3kT} \tag{47.4}$$

where N is the number of atoms per unit volume. In equation (47.4) the fact that $\mu = \mu_0(1 + \chi_p)$ is taken into account. However, since $\chi_p \ll 1$ for the majority of paramagnetic substances, we may ignore χ_p, compared with unity, in the denominator of (47.4).

Equation (47.4) shows that the paramagnetic susceptibility for constant volume (N = const) is inversely proportional to absolute temperature. This is called the *Curie law.* There is good experimental support for it in the case of paramagnetic gases and a number of paramagnetic solids. Since equation (47.4) was deduced on the assumption that $M_a B \ll kT$, deviations from the Curie law will be observed in very strong fields and at very low temperatures.

In sufficiently strong fields, such that an equation of the (42.22) type applies, we have, in the present case

$$\langle M_{az}\rangle = M_a \tag{47.5}$$

i.e., saturation is observed. All the magnetic moments are oriented along the field, and a further increase of the field produces no further increase in the magnetization. In this case

$$I_{\text{sat}} = M_a N \tag{47.6}$$

The Curie law does not hold for many paramagnetic liquids and solids, to which the elementary theory we have described is inapplicable, since it does not take into account all the factors affecting the interaction between atoms. We shall not, however, discuss the refinements of the theory necessary in these cases.

§48. Remarks on Ferromagnetism

Substances are said to be *ferromagnetic* if their permeability μ is much greater than μ_0 and depends in a complicated manner on **H**. Ferromagnetic substances possess the property of residual magnetism, the magnetization being nonzero in the absence of an external magnetic field.

The theory of ferromagnetism cannot be developed rigorously within the framework of classical electrodynamics; it is necessary to use the assumptions of quantum mechanics. Ferromagnetism is caused by the spin magnetism of electrons.

As quantum theory and experiments show, the electron possesses an internal magnetic moment and angular momentum. The internal angular momentum of the electron is called the *spin*. The internal magnetic moment and angular momentum of the electron cannot be represented by rotation of the electron about an axis, since in this case, with reasonable assumptions about the dimensions of the electron, it would be necessary to allow the linear velocity of this rotation to exceed the speed of light, which is impossible (see Part III). Therefore, the presence of the internal magnetic moment and angular momentum cannot be given a mechanical interpretation. The internal magnetic moment $\mathbf{M}_s$ of the electron, and its internal angular momentum, l_s, are related by

$$\mathbf{M}_s = \frac{e}{m}\,\mathbf{l}_s \tag{48.1}$$

Comparing this equation with equation (45.32), we see that the ratio of the internal magnetic moment to the internal angular momentum of the electron, obtained from (48.1)

$$\frac{M_s}{l_s} = \frac{e}{m} \tag{48.2}$$

is twice as large as the ratio of the magnetic moment to the angular momentum for the orbital motion of the electron, obtained from equation (45.32)

$$\frac{M}{L} = \frac{e}{2m} = \frac{1}{2}\frac{M_s}{l_s} \tag{48.3}$$

The magnetic moment of an atom is composed of the magnetic moments due to the orbital motion of the electrons (45.32) and the internal magnetic moments of the electrons (48.1). The magnetic moment of an atom $\mathbf{M}_a$ is related to its angular momentum by

$$\mathbf{M}_a = \gamma \mathbf{L}_a \tag{48.4}$$

where the value of γ, generally speaking, lies between e/m and e/2m, depending on the relative contributions of the internal and orbital moments of the electron to the total magnetic moment.

As quantum theory and experiments show, ferromagnetism is caused by the internal, or spin, magnetism of the electrons. Due to the so-called exchange forces, which are associated with the quantum-mechanical nature of the interaction, the spins of the electrons in neighboring atoms tend, under certain conditions, to become parallel to each other. This orientation of the electron spins makes a substance magnetized, and is the cause of ferromagnetism. Since the spin magnetism of the electrons, and the exchange mechanism, cannot be explained within the framework of classical electrodynamics, the theory of ferromagnetism cannot be discussed more fully here, and we must confine ourselves to these brief remarks. Experimental proof of the fact that ferromagnetism is indeed caused by spin magnetism will be given in the next section.

§49. Gyromagnetic Effects

Magnetization of a magnetic material orients the magnetic moments of individual atoms along some preferred direction. Since the magnetic moment of an atom is related to its angular momentum by (48.4), this orientation of the magnetic moments is accompanied by a corresponding orientation of the angular momenta.

The Einstein-de Haas Experiment. Let us consider a rod made of some magnetic material, suspended by an elastic thread (Fig. 64). The magnetization of this rod means that every atom acquires a magnetic moment along the direction of magnetization, which is accompanied by a corresponding change in the angular momentum of the atoms. Hence, if the

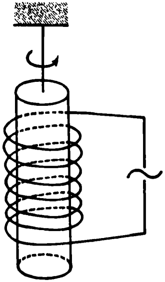


Fig. 64

rod is magnetized along the axis, the angular momentum of the rotation of the atoms about the rod axis also changes. The total angular momentum of the rod consists of the angular momentum of individual atoms and the angular momentum of the rod as a whole. Before magnetization, the total angular momentum of the rod is zero. In an isolated system, the total angular momentum is conserved. In our case, the isolated system consists of the rod and the magnetizing electromagnetic field. The magnetic field vector **H** is directed along the rod axis, while **E**, the vector of the induced electric field, is tangential to concentric circles with their centers on the axis of the rod. Thus, the Poynting vector **E** × **H** lies on a line passing through the axis of the rod and, consequently, the electromagnetic field momentum also lies along a line passing through this axis. Hence, the electromagnetic field angular momentum about the axis of the rod is zero, and we may conclude that the theorem of the conservation of the angular momentum of an isolated system reduces, in this case, to the conservation of the total angular momentum of the rod. Therefore, after the magnetization, the angular momenta of the rod is still equal to zero. But the angular momenta of the atoms change as a result of the magnetization, and, hence, the angular momentum of the rod as a whole must change by an equal and opposite amount. Thus, as a result of the magnetization, the rod begins to rotate as a whole, and twists the thread suspending it.

Let us take the z axis along the axis of the rod. From (48.4) it follows that

$$I = \sum_i \Delta M_{nzi} = \gamma \sum_i \Delta L_{nzi} = -\gamma Q_z \tag{49.1}$$

where Q_z is the angular momentum of the rod as a whole due to the magnetization, and I is the intensity of magnetization of the whole rod. The summation is performed over all molecules in the rod. The angular velocity of the rod is related to the angular momentum Q_z by the equation

$$Q_z = J\omega \tag{49.2}$$

where J is the moment of inertia of the rod. The kinetic energy of rotation is equal to

$$W_K = \frac{1}{2} J\omega^2 \tag{49.3}$$

On the other hand, the torsion modulus of the thread d is related to the frequency of the free angular oscillations ω_0 of the rod by the expression

$$J\omega_0^2 = d \tag{49.4}$$

As a result of acquiring the kinetic energy (49.3), the rod twists the thread through an angle Θ defined by the kinetic energy (49.3) and the potential energy of the thread

$$\frac{1}{2} J\omega^2 = \frac{1}{2} d\Theta^2 \tag{49.5}$$

From (49.5), using (49.4), (49.2), and (49.1), we obtain

$$J\omega = \sqrt{d}\Theta \frac{\sqrt{d}}{\omega_0} = Q_z = -\frac{1}{\gamma} I \tag{49.6}$$

Hence, it follows that

$$\gamma = -\frac{I\omega_0}{\Theta d} \tag{49.7}$$

All the quantities on the right-hand side of this equation are, in principle, measurable. This makes it possible to measure γ.

The effect is fairly small. Hence, in practice, the experiment is not carried out for a single magnetization, as described above, but using repeated magnetization reversal of a specimen at a resonance frequency ω_0. This produces a gradual increase of the torsional oscillations of the rod. However, this method introduces no changes in principle into the discussion we have given.

The Einstein-de Haas experiments were carried out on ferromagnetic rods in which the magnetization effect was especially noticeable. Within the limits of experimental error, the following value for γ was obtained

$$\gamma = \frac{e}{m} \tag{49.8}$$

Thus, a value was obtained, twice as large as would be expected if the magnetism were caused only by the orbital motion of the electrons in

atoms. At the time of these experiments, the result (49.8) was not intelligible. Later, it was shown that this result is a confirmation of the spin nature of ferromagnetism, as is evident from a comparison of equations (45.32), (48.1) and (48.4), (49.8).

Barnett Effect. When a magnetic substance is placed in a magnetic field, the magnetic moments of atoms begin to precess about the direction of the magnetic field, and the magnetic substance becomes magnetized. Thus, the magnetization is caused by the ordered precession of all the atoms in a magnetic substance with respect to the substance as a whole. Let us make the magnetic substance rotate as a whole. Individual atoms will then act as small gyroscopes, tending to retain the direction of their axes of rotation in space. Hence, the direction of the magnetic moments of individual atoms in space will not change, and these magnetic moments will perform an ordered precession about the axis of rotation of the substance, the frequency of the precession being the frequency of rotation of the magnetic substance as a whole. But this ordered precession of the atoms with respect to the substance as a whole leads to the magnetization. Hence, as a result of such rotation, the substance becomes magnetized. This effect was first observed by Barnett in 1909.

It is evident from the above discussion that if a magnetic substance rotates with a frequency ω, its magnetization is the same as it would experience if placed in a magnetic field

$$B = \frac{2m}{|e|}\omega \tag{49.9}$$

PROBLEMS

1 The diamagnetic susceptibility of copper (in the solid phase) is $\chi_d = 8.8 \times 10^{-8}$. Determine the mean distance of the electrons from the nucleus in an atom of copper.
Hint: Use equation (46.12)

$$\textit{Answer:}\quad \sqrt{\langle R^2\rangle} = \left(\frac{6m\chi_d}{e^2 Z\mu_0 N}\right)^{1/2} = 0.9 \times 10^{-10}\,\text{m}$$

2 The magnetic moment of the oxygen molecule is $M = 2.6 \times 10^{-23}$ amp m². Determine the paramagnetic susceptibility of oxygen under normal conditions.
Hint: Use equation (47.4)

$$\textit{Answer:}\quad \chi_p = \frac{M^2 N\mu_0}{3kT} \approx 18 \times 10^{-7}$$

3 Calculate the magnitude of the magnetization vector for gaseous oxygen in a weak magnetic field $H = 50$ Oe.

$$\textit{Answer:}\quad I = \chi_p H = 7.1 \times 10^{-3}\,\text{amp/m}$$

CHAPTER 10

Conductors

§50. Electrical Conductivity of Gases

Self-Maintaining and Nonself-Maintaining Currents. An electric current represents the motion of charges, and, hence, it is always associated with the motion of some charge carriers or other. The charge carriers may be electrons, ions or heavy charged particles.

A gas containing no charged particles cannot conduct electricity. A gas becomes a conductor only in the ionized state, i.e., when electric charge carriers appear in it in the form of free electrons and ions. Positive ions are atoms, molecules or groups of molecules which have lost one or more electrons. The ions may be singly or multiply charged, according to the number of electrons they have lost. Negative ions are atoms, molecules, or groups of molecules which have gained electrons. Negative ions are usually singly charged.

Under normal conditions, all gases are poor conductors of electricity. They only become good conductors in the presence of some external factor producing ionization (high temperature, irradiation with ultraviolet or x-rays, etc). If the electric field is fairly weak, then the current through a gas ceases when the action of the external ionization factor ceases. Such currents are said to be *nonself-maintaining.*

If the electric field is sufficiently strong, then it may itself produce ionization in the gas, so that the gas becomes a conductor. The current arising in this case is called a *self-maintaining* current. In this case there is no general function giving the dependence of the current on the applied field. Everything depends on the actual conditions. In particular, quite

often the self-maintaining current decreases as the applied field increases.

Nonself-Maintaining Current. Let us consider the case of a nonself-maintaining current in greater detail. Let N denote the number of ions of given sign per unit volume. Let the rate of formation of ions by an external source be dN/dt per unit volume. The number of ions is reduced by recombination, i.e., mutual neutralization of the ions. After a sufficient period of time, when the processes of ionization and recombination reach a state of equilibrium, we may assume that the numbers of negative and positive ions are equal

$$N^{(+)} = N^{(-)} = N \tag{50.1}$$

where the ions are taken to be singly charged.

Clearly, the rate of recombination is proportional to the product of the concentrations of the ions, i.e., to N^2. Hence, at equilibrium, we have the equation

$$\left(\frac{dN}{dt}\right) = -rN^2 \tag{50.2}$$

where r is called the *coefficient of recombination.*

The current density is, evidently

$$j = j^{(+)} + j^{(-)} = e(N^{(+)}v^{(+)} + N^{(-)}v^{(-)}) = eN(v^{(+)} + v^{(-)}) \tag{50.3}$$

where $v^{(+)}$ and $v^{(-)}$ are the velocities of the ordered drift motion of the positive and negative ions, respectively, under the action of the applied electric force. The drift velocity of an ion in an electric field is proportional to the field

$$v = gE \tag{50.4}$$

where g is called the *mobility of the ion.* It is numerically equal to the drift velocity in an electric field of intensity $E = 1$ V/m. The mobilities of the positive and negative ions $g^{(+)}$ and $g^{(-)}$ are, generally speaking, different. Using (50.4), (50.3) may be rewritten

$$j = e(g^{(+)} + g^{(-)})NE \tag{50.5}$$

The form of this equation recalls Ohm's law. However, it is equivalent to Ohm's law only when the coefficient of E is independent of E and j. In general, this coefficient does depend on E and j, and, hence, the resemblance between (50.5) and Ohm's law is merely formal.

When the number of gaseous ions recombining per unit time is much greater than the number of ions impinging on an electrode per unit time, we may use the equilibrium value for N from (50.2). Substituting it in (50.5), we obtain

$$j = e(g^{(+)} + g^{(-)})\sqrt{\frac{1}{r}\frac{dN}{dt}}\,E \tag{50.6}$$

To find the conditions under which this expression is valid, we must bear in mind that the mobilities of ions at normal pressure are of the order of 1 cm/sec per 1 V/cm and the recombination coefficients are $r \approx 10^{-6}$ cm^3/sec. For example, if dN/dt is of the order of 10^{10} ions per cm^3 per sec, and $E = 10$ V/m, the number of ions impinging on 1 cm^2 of the electrode per unit time equals

$$\frac{j}{e} = (g^{(+)} + g^{(-)}) \left(\frac{1}{r}\frac{dN}{dt}\right)^{1/2} E \approx 2 \times 10^9 \tag{50.7}$$

If the distance between a pair of plane electrodes is 10 cm, then 10^{11} ions recombine per second per 1 cm^2 between the electrodes, i.e., the condition of applicability of (50.6) is satisfied because the rate of recombination is much greater than the rate at which ions reach the electrodes.

Saturation Current. Let the distance between a pair of plane electrodes be d. If the electric field is so strong that all the ions formed by an external source reach the electrodes before they have a chance to recombine, then we have a saturation current. The saturation current, is clearly equal to

$$j_{\text{sat}} = ed \left(\frac{dN}{dt}\right)_{\text{gen}} \tag{50.8}$$

Characteristic of a Discharge. In moderate fields some of the ions recombine before reaching the electrodes. The balance between the loss and the generation of the ions is written in the form

$$\left(\frac{dN}{dt}\right)_{\text{gen}} + \left(\frac{dN}{dt}\right)_{\text{rec}} + \left(\frac{dN}{dt}\right)_{j} = 0 \tag{50.9}$$

Hence, using (50.8), (50.2), and (50.3), we obtain

$$\frac{j_{\text{sat}}}{e} - rN^2d - N(g^{(+)} + g^{(-)})E = 0 \tag{50.10}$$

But since

$$j = eN(g^{(+)} + g^{(-)})E \tag{50.11}$$

we can rewrite (50.10) as an equation for j

$$j^2 + 2\alpha j - 2\alpha j_{\text{sat}} = 0 \tag{50.12}$$

where

$$\alpha = \frac{|e|(g^{(+)} + g^{(-)})^2E^2}{2rd} \tag{50.13}$$

The positive root of this quadratic equation has the form

$$j = \alpha\left(\sqrt{1 + \frac{2j_{\text{sat}}}{\alpha}} - 1\right) \tag{50.14}$$

The graph of the current density is shown in Fig. 65. In the limiting cases, $\alpha \ll j_{sat}$ and $\alpha \gg j_{sat}$, this solution is transformed into equations (50.6) and (50.8).

The expression (50.14) is called the *characteristic* of the discharge. It agrees well with experiment if the additional loss of ions due to diffusion is taken into account.

Self-Maintaining Current. If we continue to increase the intensity of the electric field after the saturation current is reached, then, at some value of the field, the current starts to increase again. This is due to the fact that the electrons in the gas that have not yet become attached to molecules can be accelerated (by the strong field) to energies at which they themselves can bring about ionization of the gas molecules. As a result, the rate of ionization becomes dependent on the intensity of the applied field. The current thus produced is called the self-maintaining current. In Fig. 65,

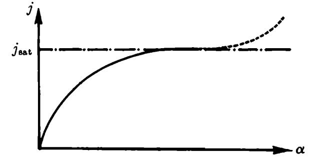


Fig. 65

the initial portion of the characteristic of this current is represented by the dotted curve.

Effect of a Space Charge. As we have already pointed out, the mobilities of the positive and negative ions are, generally speaking, different, and usually $g^{(-)} > g^{(+)}$, so that the current density due to the motion of the positive ions is less than the current density due to the motion of the negative charges. Therefore, the number of the positive ions impinging on the cathode is less than the number of negative charges striking the anode, although the number of ions being formed and recombining is the same for both. It is clear that such a state cannot be in equilibrium. An equilibrium state is reached as follows. As a result of the motion of the positive charges to the cathode, and the negative charges to the anode, the cathode acquires an excess of the positive charges and the anode an excess of the negative charges. However, due to the higher mobility of the negative charges, the excess of the negative charge on the anode will be less than the excess of the positive charge on the cathode. As a result of

this redistribution of the charges and the accompanying changes in the electric field, a state of equilibrium is reached in which the number of positive and negative ions striking the respective electrodes becomes the same.

Mobility of the Ions. An ion of mass M and charge e in a homogeneous field E moves with a constant acceleration

$$a = \frac{eE}{M} \tag{50.15}$$

and, if the initial velocity of the ion is zero, it moves in the direction of the field and travels a distance

$$s = \frac{1}{2}\frac{eE}{M}\tau^2 \tag{50.16}$$

in time τ. If the mean free path of an ion in a gas is l, and the mean velocity is $\langle v\rangle$, then we may assume

$$\tau = \frac{l}{\langle v\rangle} \tag{50.17}$$

Assuming that after every collision the ion loses all its energy of ordered motion, we may, using equations (50.16) and (50.17), write down the following expression for the drift velocity

$$v_d = \frac{1}{2}\frac{eE}{M}\tau = \frac{1}{2}\frac{el}{M\langle v\rangle}E \tag{50.18}$$

The refinements brought about by taking the statistical distribution of the mean free path into account involve only a slight change in the numerical coefficient in equation (50.18). Hence, we can write down the expression for the mobility of the ions

$$g = \frac{1}{2}\frac{el}{M\langle v\rangle} \tag{50.19}$$

From this equation, it is clear that the mobility of positive and negative ions should be the same. This is actually true only if the electron component in the value of the mobility of the negative ions is small. Otherwise, due to the transport of charge by the motion of electrons, the mobility of the negative ions is higher. It is clear from equation (50.19) that the mobility should be inversely proportional to the density of the gas, since the mean free path l is inversely proportional to the density. This is confirmed by experiment.

However, on the whole, equation (50.19) does not explain the experimental facts completely. In particular, experiment does not confirm the corollary, deduced from this equation, that the mobility should be propor-

tional to the square root of temperature (since $l \sim T$, $\langle v \rangle \sim \sqrt{T}$). Experiment actually gives mobility values several times smaller than the theoretical values. In order to explain this difference, Langevin studied the polarization of ions as they approach each other during a collision. He found that the ions acquire dipole moments, and this alters the nature of the collision. To take this phenomenon into account, substantial changes must be made in the equations. It is not possible to describe this theory in the present book.

§51. Electrical Conductivity of Liquids

The majority of pure liquids are bad conductors of electricity, i.e., they contain very few charge carriers. Solutions of salts, acids, and bases in water, and in some other liquids, are good conductors of electricity, since the molecules of the solute dissociate into positive and negative ions. It is these ions which are responsible for the conductivity of the solution. If there is no dissociation into ions, then the solution is not a conductor of electricity.

Denoting by $N = N^{(+)} = N^{(-)}$ the number of ions of one sign in the solution, we may write down the expression for the current density

$$j = e(g^{(+)} + g^{(-)})NE \tag{51.1}$$

where e is the charge of the ions, $g^{(+)}$ and $g^{(-)}$ are the mobilities of the positive and negative ions, which are defined in exactly the same way as for the ions in gases. The mobility of ions in liquids is much lower than the mobility of ions in gases, and is of the order of 0.001 cm/sec per 1 V/cm.

The concentration of the ions depends on the degree of dissociation. The latter is described by a dissociation coefficient α, which is defined as the ratio of the concentration of the dissociated molecules N to the concentration of the solvent molecules N_0

$$N = \alpha N_0 \tag{51.2}$$

Hence, the concentration of nondissociated molecules is equal to

$$N' = (1 - \alpha)N_0 \tag{51.3}$$

In a solution we have a continuous process of the dissociation of molecules and the recombination of ions to form neutral molecules. In the equilibrium state the two processes balance each other out. The number of dissociating molecules is, clearly, proportional to the number of nondissociated molecules (51.13), i.e.

$$\Delta N = \beta(1 - \alpha)N_0 \tag{51.4}$$

where β is a coefficient of proportionality. The number of recombining molecules $\Delta N'$ is proportional to the product of the concentrations of the positive and negative ions, i.e.

$$\Delta N' = \gamma\alpha^2 N_0^2 \tag{51.5}$$

where γ is a coefficient of proportionality. In the equilibrium state we have

$$\Delta N = \Delta N' \tag{51.6}$$

Hence, using (51.4) and (51.5) we obtain the following expression, which relates the degree of dissociation to the concentration of the solute

$$\frac{1-\alpha}{\alpha^2} = \frac{\gamma}{\beta} N_0 \tag{51.7}$$

It is clear that the dissociation coefficient depends on the concentration of the solute. At very weak concentrations, when $N_0 \approx 0$, we have, from (51.7)

$$\alpha \approx 1 \tag{51.8}$$

i.e., the dissociation is almost complete.

If there is little dissociation, i.e., when $\alpha \ll 1$, we obtain, from (51.7)

$$\alpha = \sqrt{\frac{\beta}{\gamma}} \frac{1}{\sqrt{N_0}} \tag{51.9}$$

i.e., the dissociation coefficient decreases as the concentration of the solute increases.

Using (51.2), equation (51.1) may be written

$$j = e(g^{(+)} + g^{(-)})\alpha N_0 E \tag{51.10}$$

The motion of ions in a liquid may be considered, over a wide range of conditions, as the motion in a viscous medium. For such motion, the velocity is proportional to the applied force. In the present case, the applied force is proportional to the intensity of the field. Hence, the velocity of ions is proportional to the electric field

$$v = gE \tag{51.11}$$

where g is the mobility which, as has been shown, is independent of the field intensity. It is only in very strong fields, of the order of megavolts per centimeter, that there is a deviation from the direct proportionality between the field and the ion velocity. The dissociation coefficient in (51.10) is also independent of E over a wide range of fields. Hence, right up to very strong electric fields in the megavolt per centimeter range, equation (51.10) expresses Ohm's law. Hence, we can write down the following expression for the electrical conductivity of the solution

$$\lambda = e(g^{(+)} + g^{(-)})\alpha N_0 \tag{51.12}$$

At low concentrations, the dissociation coefficient, defined by (51.8), is constant, and the sum of the mobilities of the ions $g^{(+)} + g^{(-)}$ is also approximately constant. Hence, at low concentrations, the conductivity is proportional to the concentration. At high concentrations, the dependence of the conductivity on the concentration becomes considerably more complicated. On the one hand, we must take into account the dependence of the dissociation coefficient on the concentration, equations (51.7), (51.9), and on the other hand, the mobility of the ions begins to depend markedly on the concentration, decreasing in concentrated solutions, since the electrical interaction between the ions begins to play a part in the process. Hence, at high concentrations no direct proportionality between the conductivity and the concentration is observed.

§52. Electrical Conductivity of Metals

The electrical conductivity of metal conductors is due to the motion of free electrons which are present in these conductors in the form of the "electron gas." The presence of free electrons in metal conductors can be demonstrated by mechanical tests.

Tolman and Stewart's Experiment. The apparatus consists of a coil of wire which can rotate about its axis. The ends of the coil are connected to a galvanometer by sliding contacts (Fig. 66). If the coil is set in rapid rotation and then suddenly stopped, free electrons in the wire will con-

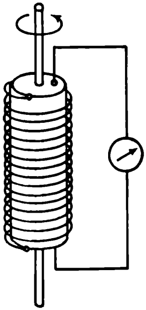


Fig. 66

tinue moving forward under their inertia, and the galvanometer will register a current pulse.

Let us denote the linear deceleration during the stopping state by $\dot{v}$. This deceleration acts along the tangent to the circumference of the coil. If the wire is sufficiently thin, and sufficiently densely wound, then we may assume that this deceleration acts along the wire, and that an inertial force $-m\dot{v}$ acts, opposite to the direction of $\dot{v}$, on every free electron. Under the action of this force, each free electron in the coil wire behaves as if there were some effective electric field E_{eff} in the wire

$$E_{\text{eff}} = -\frac{m}{e}\dot{v} \tag{52.1}$$

Thus, the effective emf in the coil, caused by the inertia of free electrons, is equal to

$$\mathcal{E}_{\text{eff}} = \int_L E_{\text{eff}}\, dl = -\frac{m}{e}\dot{v}\int_L dl = -\frac{m}{e}\dot{v}L \tag{52.2}$$

where L is the length of the conductor wound into the coil. All parts of the coil are retarded by the same deceleration, hence, $\dot{v}$ can be taken outside the integral sign in (52.2).

We shall denote the current passing round the closed circuit by I, and the resistance of the whole circuit by R. Hence, we may write down Ohm's law in the form

$$IR = -\frac{m}{e}\dot{v}L \tag{52.3}$$

The quantity of electricity passing across the cross section of the conductor in time dt, when the current is I, is equal to

$$dq = I\, dt = -\frac{m}{e}\frac{L}{R}\dot{v}\, dt = -\frac{m}{e}\frac{L}{R}\, dv \tag{52.4}$$

Hence, during the time interval needed to slow down the coil from its initial linear velocity v_0 to complete standstill, the quantity of electricity passing through the galvanometer is

$$q = \int dq = -\frac{m}{e}\frac{L}{R}\int_{v_0}^{0} dv = \frac{m}{e}\frac{L}{R}v_0 \tag{52.5}$$

The value of q is found from the galvanometer reading; L, R, and v_0 are known. Hence, we can find both the sign and the value of e/m. Experiments show that the current observed with a galvanometer is, indeed, due to the motion of electrons.

Elementary Theory of Electrical Conductivity. Let us denote the conduction electron density in a metal by N. The mechanism of conduction in a

metal may be considered, in the first approximation, to resemble the process in gases, except that instead of ions, the charge carriers are electrons. Under the action of an electric field, the conduction electrons are accelerated, lose their velocity by collisions with atoms, are again accelerated, etc. If we repeat the argument by which equation (50.18) was deduced, we obtain the following equation for the drift of electrons in a metal

$$v_d = \frac{1}{2}\frac{el}{m\langle v\rangle}E \tag{52.6}$$

where m, $\langle v\rangle$, and l are, respectively, the mass of one electron, the mean velocity of the disorderly motion of the electrons, and the electron's mean free path in the case of collisions with the atoms of the metal. Hence, we obtain the following expression for the current density

$$j = ev_dN = \frac{1}{2}\frac{e^2l}{m\langle v\rangle}NE \tag{52.7}$$

Comparing (52.7) with the differential form of Ohm's law, we see that the electrical conductivity of a metal conductor is equal to

$$\lambda = \frac{1}{2}\frac{e^2l}{m\langle v\rangle}N \tag{52.8}$$

Joule-Lenz Law. This representation of the mechanism of conductivity also allows us to obtain a correct expression for the Joule-Lenz law.

At the end of a free path, the velocity of an electron, according to (50.15), is equal to

$$v_{\text{final}} = a\tau = \frac{eE}{m}\frac{l}{\langle v\rangle} \tag{52.9}$$

Hence, when an electron collides with an atom in a metal, it transfers to the atom its own kinetic energy, equal to

$$W_K = \frac{mv_{\text{final}}^2}{2} = \frac{e^2E^2l^2}{2m\langle v\rangle^2} \tag{52.10}$$

The mean number of collisions per unit time is $\langle v\rangle/l$. Therefore, the N electrons per unit volume transfer an energy $\bar{q}$ to the metal per unit time, where

$$\bar{q} = W_K\frac{\langle v\rangle}{l}N = \frac{1}{2}\frac{e^2l}{m\langle v\rangle}NE^2 \tag{52.11}$$

Equation (52.8) shows that this equation expresses the Joule-Lenz law

$$\bar{q} = \lambda E^2 \tag{52.12}$$

Insufficiency of the Classical Theory of Electrical Conductivity. In spite of the obvious and natural nature of this elementary theory of electrical

conductivity, it cannot give correct quantitative results. Experiment shows that the electrical conductivity is inversely proportional to absolute temperature ($\lambda \propto T^{-1}$). It is impossible to explain this relationship on the basis of equation (52.8), since $\langle v \rangle \sim \sqrt{T}$, and it is not possible to assume that $lN \sim 1/\sqrt{T}$.

Another contradiction appears in connection with the specific heat of conductors. We shall assume that a conductor contains many free electrons which move as if in a gas. From the theorem on the uniform distribution of energy over the degrees of freedom, we deduce that the mean kinetic energy of electrons in a metal should be equal to the mean kinetic energy of molecules of the metal. To obtain reasonable values of the conductivity, we have to assume that the number of free electrons is approximately equal to the number of atoms in the metal. Hence, the specific heat of conductors must be considerably higher than the specific heat of dielectrics, due to the presence of free electrons. However, this is not observed, and we therefore reach the conclusion that the conduction electrons take part in the process of conducting electricity, but do not affect the specific heat of conductors.

Another fact is also difficult to understand in terms of the classical presentation: to make equation (52.8) agree with experiment, we have to make the mean free path l very large, some thousands of times greater than the distance between the atoms, which is inconceivable.

All these contradictions are of a basic nature and they cannot be overcome merely by improvements of the classical theory.

Quantum theory gives a completely different treatment of the phenomenon of the conduction of electricity from the point of view of the band theory of solids. This can be found in a treatise on quantum theory of solids, or quantum statistical mechanics.

Methods of Calculating the Resistance of a Medium to an Electric Current. If a constant current flows along a thin conductor, then, if we know the conductivity of the material of the conductor, it is easy to find the resistance R, and to write down Ohm's law for the conductor

$$RI = U \tag{52.13}$$

where I is the current in the conductor, and U is the potential difference between the electrodes to which the conductor is connected. In the case of a thin conductor

$$R = \frac{1}{\lambda}\frac{l}{S} \tag{52.14}$$

where λ, l and S are the conductivity, length and cross sectional area of the conductor, respectively.

This simple method is not applicable, however, to currents flowing in massive conductors of composite form, or in conducting media. For example, let us suppose that the electrodes have fallen on to the ground, and that the conductivity of the soil is known. What will be the resistance of the soil to the current between the electrodes? Another example: consider electrodes in the form of two coaxial cylinders of different radii, the space between them being filled with a conducting medium of conductivity λ. Find the resistance of the conducting medium in the cylindrical region between the electrodes. There exists a general method of solving problems of this type, as follows.

By definition, the resistance of a medium between two electrodes is the ratio of the potential difference between the electrodes to the current flowing between them. The current lines and the current density may be found using the differential form of Ohm's law

$$\mathbf{j} = \lambda \mathbf{E}$$

In general, the first step is to determine the electric field set up by electrodes of known form and at known potentials. We have already considered the solution of this type of problem. Having determined the form of the lines of force of the field, we automatically determine the current lines. The current density at every point is given by: $\mathbf{j} = \lambda\mathbf{E}$. Thus, the problem of the passage of the current is completely solved. However, in the majority of cases there is no need to know the form of the current lines and current density at all points; it is sufficient to know only the ratio of the potential difference between the electrodes and the current flowing between them, i.e., it is sufficient to know the resistance. In this case, it is useful to know the relationship between the resistance and the capacitance which would exist between the electrodes if the conducting medium were removed. This relationship may be found as follows. The current flowing from one electrode to the other is

$$I = \int_{S'} \mathbf{j}\cdot d\mathbf{S} = \lambda \int_{S'} \mathbf{E}\cdot d\mathbf{S} \tag{52.15}$$

where S' is an almost closed surface surrounding an electrode, except for a small area through which the current flows to the electrode. If the electrode is considered as an insulated conductor at a definite potential, then by Gauss' theorem

$$\oint_S \mathbf{E}\cdot d\mathbf{S} = \frac{q}{\epsilon_0} \tag{52.16}$$

where S is a closed surface surrounding the electrode, and q is the charge on the electrode. Since S and S' differ only by a very small area, the

integral over S' in (52.15) is very slightly different from the integral over S in (52.16), so that we may write

$$I = \frac{\lambda}{\epsilon_0} q \tag{52.17}$$

The electrostatic capacitance between two electrodes, considered as plates of a capacitor, is $C = q/U$, where U is the potential difference between the plates. Equation (52.17), therefore, may be put in the form

$$I = \frac{\lambda}{\epsilon_0} CU \tag{52.18}$$

Using the definition of the resistance (52.13), we obtain the following expression for the resistance between the electrodes

$$R = \frac{\epsilon_0}{\lambda C} \tag{52.19}$$

Thus, if the capacitance between two electrodes and the conductivity are known, then the resistance between the electrodes may be calculated from equation (51.19).

As an example, we shall determine the resistance of a medium between two coaxial cylindrical electrodes, radii r_1 and r_2. The conductivity of the medium is λ, and the length of the cylindrical electrodes is l. The capacitance of a cylindrical capacitor of these dimensions is given by

$$C = \frac{2\pi\epsilon_0 l}{\ln \frac{r_2}{r_1}}$$

Hence, according to equation (52.19), the resistance between the cylindrical electrodes is

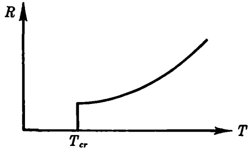
$$R = \frac{\ln \frac{r_2}{r_1}}{2\pi l \lambda}$$

§53. Superconductivity

The phenomenon of superconductivity was discovered by Kammerlingh-Onnes in 1911. At very low temperatures, of the order of several degrees Kelvin, certain pure metals lose their electrical resistance. Some alloys are superconducting as well. In fact, the transition temperature can be altered by alloying. It must be stressed that we are discussing the complete loss of resistance, and not merely a considerable reduction. Transition to the superconducting state occurs suddenly, at some critical temperature, so that the temperature dependence of the resistance of a superconductor is

of the form shown in Fig. 67. Certain alloys also possess superconducting properties. Among pure metals which exhibit superconductivity are aluminum, mercury, lead, tin, zinc, and several others. However, superconductiv-

Fig. 67



ity is not a universal property; such metals as copper, silver, gold and others do not become superconductors even at temperatures of the order of 0.1° K.

The mechanism of superconductivity remained a mystery for a very long time, and has been explained only in the last few years. It is connected with the quantum properties of the motion of electrons in metals. Here we shall confine ourselves to a description of the basic experimental facts connected with superconductivity.

Although the electrical resistance of a superconductor is zero, it is impossible to consider it as an ideal conductor with zero resistance. A superconductor differs from such an ideal conductor primarily because there is no magnetic field within it: $\mathbf{B} = 0$. Hence, using the boundary condition (9.10), we may conclude that on the outer surface of a superconductor the normal component of the magnetic induction is also zero. This means that the magnetic lines of force envelop the superconductor, running parallel to its surface.

From $\mathbf{B} = 0$, it follows that inside a superconductor the tangential component of the magnetic field is zero. Outside a superconductor, the magnetic field is not equal to zero. Therefore, there must be a discontinuity in the tangential component of the magnetic field on the surface of a superconductor. In accordance with the boundary condition (9.30), this is possible only in the case of surface currents. Hence, we may conclude that the current in a superconductor flows along the surface. From (9.30), it is clear that the surface current density i_{surf} is related to the magnetic field H, close to the surface of the superconductor, by the expression

$$i_{\text{surf}} = H \tag{53.1}$$

where i_{surf} flows in a direction perpendicular to the direction of $\mathbf{H}$, as follows from the derivation of equation (9.30). Surface currents also appear when

bodies are magnetized. However, these surface currents always cancel each other out, so that there is no net current. In superconductors, however, the presence of surface currents does lead to a nonzero net current. Hence, if a superconductor forms a part of a closed circuit, a current may flow in it, although a source of emf must be included in the circuit in order to maintain the current in the nonsuperconducting part of the circuit. If the superconductor makes up the whole closed circuit, e.g., if it is in the form of a ring, then a steady superconducting current flows in it for as long as may be desired, without needing any constant emf to maintain it. This means that the current flows in the superconductor without any loss of energy and there is no need to make up for this loss. This property of a superconductor is also described as the absence of the electrical resistance.

The transition from a state of zero magnetic field inside a superconductor to a finite magnetic field outside it does not take place suddenly, but step-by-step, although very rapidly. The magnetic field does, in fact, penetrate the superconductor, to a depth of the order of 10^{-5} cm, which, in comparison with interatomic distances (10^{-8} to 10^{-7} cm) is quite large. Hence, we can speak about only the absence of the magnetic field in superconductors with sufficient accuracy when we are discussing massive superconductors. In sufficiently thin superconductors (films, thin wires), the magnetic field penetrates almost the whole thickness of the superconductor, and is somewhat weaker in the center of the superconductor.

The superconducting state cannot exist in all magnetic fields. If there is a sufficiently strong magnetic field close to the surface, the superconductivity is destroyed. This value of the magnetic field H_{cr} is called the *critical field.* For the majority of massive superconductors, the critical magnetic field is of the order of some hundreds of oersteds. When the current flowing in a superconductor sets up a critical magnetic field close to its surface, the superconductivity is likewise destroyed. Thus, there is a critical value of the current flowing in a superconductor. The relationship between the critical current I_{cr} and the critical value of the magnetic field H_{cr} for a massive cylindrical conductor may be found from (53.1). The total current flowing along the surface of a cylindrical superconductor of radius r is

$$I = 2\pi r i_{surf}$$

Hence, using (53.1), it follows that

$$I_{cr} = 2\pi r H_{cr}$$

For thin wires of radius r of the order of less than the thickness of penetration of the magnetic field into the superconductor, the critical field is some ten times greater than for massive conductors, and the critical current is less.

The phenomenon of superconductivity is of great practical value. Using superconductors, electric energy can be transmitted without any loss. Nowadays, static magnetic fields of some hundreds of thousands of oersteds are obtained using superconductors in solenoid form. The only other way of obtaining such fields is to use pulse conditions, but in that case the field persists for only a very short time.

PROBLEMS

1 The conduction electron density in copper is approximately 8.5×10^{22} electrons per cm^3. Determine the mean drift velocity of the conduction electrons for a current density $j = 10$ amp/mm^2

Answer: $v_{dr} = 0.0736$ cm/sec

2 A copper sphere of 10 cm diameter is dropped into a hemispherical copper cup of 20 cm diameter filled with water, in such a way that the sphere and the cup are concentric. The conductivity of the water $\lambda = 10^{-3}$ ohm^{-1} m^{-1}. Determine the electrical resistance between the sphere and the cup

Answer: $R = 1590\ \Omega$

3 A spherical electrode of radius a is placed in a medium of conductivity λ at a distance d from a very large highly conducting plate. Find the resistance to an electric current flowing from the electrode to the plate

Answer: $R = \dfrac{1}{4\pi\lambda}\left(\dfrac{1}{a} - \dfrac{1}{2d}\right)\ (a \ll d)$

CHAPTER 11

Relationship Between Phenomenological Electrodynamics and Electron Theory

§54. Averaging of Fields

The fully rigorous equations of microscopic electrodynamics (36.3) may be used to describe electromagnetic phenomena both in media and in empty space. If all charges and their motion are properly taken into account, then these equations can be used to determine the electromagnetic field in space. None of the constants which describe the properties of a medium need be introduced. These properties are automatically allowed for in the equations by the charges and currents which are present in the medium or induced in the medium by external fields. The terms of equation (36.3) are very rapidly varying functions of the coordinates and time. If we measure the intensity of the field or the current density at some point of the medium, we are actually measuring some mean values over some small volume and a short time interval. We relate these mean values to some point in the volume, and to some instant of time, by the process of averaging, and speak of these values as functions of the point and instant. These averaged quantities vary smoothly from point to point and with time. Phenomenological electrodynamics is concerned with these averaged quantities. It is clear that the equations of the phenomenological electrodynamics must be obtained if one averages the equations of the electron theory. This averaging allows the physical meaning of the quantities, on which the phenomenological electrodynamics rests, to be made clear.

Let us take as our volume, over which the averaging will be performed, a physically small volume ΔV, as defined in §42. As the time interval for averaging we choose a physically small interval Δt, defined as follows: this interval is much longer than the periods of the microscopic fields, and much shorter than the time during which the mean values vary perceptibly. The mean value of the function $f(x, y, z, t)$ is defined by

$$\langle f(x, y, z, t)\rangle = \frac{1}{\Delta V\,\Delta t}\int_{\Delta V\,\Delta t} f(x+\xi, y+\eta, z+\zeta, t+\tau)\,d\xi\,d\eta\,d\zeta\,d\tau \tag{54.1}$$

The physically small volume is chosen in the neighborhood of the point x, y, z. From the definition of this volume, it follows that its magnitude and form are, within known limits, arbitrary, and that the position of the point x, y, z, in this volume is arbitrary. The same may be said of the time interval Δt; the mean values do not change if we increase or decrease the averaging interval by a small amount. This assertion is simply another way of expressing the definition of a physically small volume and time interval.

From the definition (54.1), the following basic averaging formulas follow immediately

$$\frac{\partial}{\partial t}\langle f\rangle = \left\langle\frac{\partial f}{\partial t}\right\rangle \qquad \frac{\partial}{\partial x}\langle f\rangle = \left\langle\frac{\partial f}{\partial x}\right\rangle$$

$$\frac{\partial}{\partial y}\langle f\rangle = \left\langle\frac{\partial f}{\partial y}\right\rangle \qquad \frac{\partial}{\partial z}\langle f\rangle = \left\langle\frac{\partial f}{\partial z}\right\rangle \tag{54.2}$$

Using (54.2) we can average equations (36.3) of the electron theory. We thus obtain the following averaged equations

$$\operatorname{curl}\langle\mathbf{h}\rangle = \langle\rho\mathbf{v}\rangle + \frac{\partial}{\partial t}\langle\mathbf{d}\rangle \tag{54.3}$$

$$\operatorname{curl}\langle\mathbf{e}\rangle = -\frac{\partial}{\partial t}\langle\mathbf{b}\rangle \tag{54.4}$$

$$\operatorname{div}\langle\mathbf{b}\rangle = 0 \tag{54.5}$$

$$\operatorname{div}\langle\mathbf{d}\rangle = \langle\rho\rangle \tag{54.6}$$

$$\langle\mathbf{d}\rangle = \epsilon_0\langle\mathbf{e}\rangle \qquad \langle\mathbf{b}\rangle = \mu_0\langle\mathbf{h}\rangle \tag{54.7}$$

The corresponding forms of Maxwell's equations, written in terms of the averaged quantities, are

$$\operatorname{curl}\mathbf{H} = \mathbf{j} + \frac{\partial\mathbf{D}}{\partial t} \tag{54.8}$$

$$\operatorname{curl}\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} \tag{54.9}$$

$$\operatorname{div}\mathbf{B} = 0 \tag{54.10}$$

$$\operatorname{div} \mathbf{D} = \rho \tag{54.11}$$

$$\mathbf{D} = \epsilon \mathbf{E} \qquad \mathbf{B} = \mu \mathbf{H} \tag{54.12}$$

Equations (54.3) to (54.7) and (54.8) to (54.12) describe one and the same electromagnetic field. Hence, a comparison of these equations establishes the meaning of the fundamental physical quantities of the phenomenological electrodynamics from the point of view of the microscopic electrodynamics.

The pairs of equations (54.4) and (54.9), and (54.5) and (54.10) contain no charges or currents, and are completely identical. Comparison shows that

$$\langle \mathbf{e} \rangle = \mathbf{E} \qquad \langle \mathbf{b} \rangle = \mathbf{B} \tag{54.13}$$

The meaning of the first equation of (54.13) is clear; the electric field in the phenomenological electrodynamics is the mean value of the microscopic electric field. The meaning of the second equation of (54.13) may be demonstrated very easily if it is transformed, using equation (54.7), into

$$\mathbf{B} = \mu_0 \langle \mathbf{h} \rangle \tag{54.14}$$

Using the relationship $\mathbf{B} = \mu \mathbf{H}$ for the macroscopic field, we may rewrite (54.14) as

$$\mu' \mathbf{H} = \langle \mathbf{h} \rangle \tag{54.15}$$

where $\mu' = \mu/\mu_0$ is the relative permeability in Gaussian units. Thus, the mean value of the microscopic magnetic field is not equal to the value of the magnetic field in the phenomenological electrodynamics. Denoting the macroscopic and microscopic magnetic fields in Gaussian units by $\mathbf{H}'$ and $\mathbf{h}'$, respectively, we may rewrite (54.15) as

$$\mathbf{B}' \equiv \mu' \mathbf{H}' = \langle \mathbf{h}' \rangle \tag{54.16}$$

We sometimes say, therefore, that the induction of the macroscopic field is the mean value of the microscopic magnetic field. Of course, this assertion is contained in equation (54.14) as well; the mean value of the magnetic microscopic field leads not to the macroscopic magnetic field, but to the macroscopic magnetic induction.

§55. Averaging the Microscopic Current Density

Using (54.13) and (54.7), we can rewrite (54.3) and (54.6) as follows

$$\operatorname{curl} \frac{\mathbf{B}}{\mu_0} = \langle \rho \mathbf{v} \rangle + \frac{\partial}{\partial t} \epsilon_0 \mathbf{E} \tag{55.1}$$

$$\operatorname{div} \epsilon_0 \mathbf{E} = \langle \rho \rangle \tag{55.2}$$

These equations are considerably different from the corresponding Maxwell's equations (54.8) and (54.11). Using (14.25) and (20.22), we can put Maxwell's equations (54.8) and (54.11) in the form

$$\operatorname{curl} \frac{\mathbf{B}}{\mu_0} = \mathbf{j} + \operatorname{curl} \mathbf{I} + \frac{\partial \mathbf{P}}{\partial t} + \frac{\partial}{\partial t} \epsilon_0 \mathbf{E} \tag{55.3}$$

$$\operatorname{div} \epsilon_0 \mathbf{E} = \rho - \operatorname{div} \mathbf{P} \tag{55.4}$$

Comparison of equations (55.1) to (55.2) and (55.3) to (55.4) shows that these equations are the same when the conditions

$$\langle \rho \mathbf{v} \rangle = \mathbf{j} + \operatorname{curl} \mathbf{I} + \frac{\partial \mathbf{P}}{\partial t} \tag{55.5}$$

$$\langle \rho \rangle = \rho - \operatorname{div} \mathbf{P} \tag{55.6}$$

are satisfied. If we can prove equations (55.5) and (55.6) by a method independent of Maxwell's equations and of the equations of the electron theory, we shall prove that Maxwell's equations are obtained by averaging the equations of the electron theory, and the physical meaning of the quantities of the phenomenological electrodynamics will be explained from the point of view of the electron theory.

Matter contains neutral atoms and molecules, ions and free electrons. The microscopic current density is due to various kinds of motion of charges in matter.

Conduction Current. The motion of ions and electrons among other particles gives rise to a microscopic conduction current. In the absence of an external electric field, the velocities of individual ions and electrons have a random distribution, and, hence, the mean value of the conduction current is zero. In the presence of an external electric field, a predominant direction of motion of ions and electrons becomes evident. This ordered motion of electrons and ions produce the conduction current. The ordered motion of electrons and ions may be caused not only by an applied electric field; it may appear, for example, due to a nonuniform distribution of matter. In this case, the current flows in a direction collinear with the concentration gradient. Denoting the charge of the i^{th} particle (electron or ion) by e_i, and the velocity of the i^{th} particle by v_i, we may write down the following expression for the mean microscopic current density

$$\langle \rho \mathbf{v}_{\text{cond}} \rangle = \frac{1}{\Delta V} \sum_{\Delta V} e_i \mathbf{v}_i = \mathbf{j} \tag{55.7}$$

where ΔV is a physically small volume. The summation is performed over the particles in ΔV. Equation (55.7) represents the conduction current density $\mathbf{j}$ which occurs in (55.5).

Polarization Current. The polarization vector $\mathbf{P}$ is defined by

$$\mathbf{P} = \frac{1}{\Delta V} \sum_{\Delta V} \mathbf{p}_i \tag{55.8}$$

where $\mathbf{p}_i$ is the dipole moment of one molecule. The dipole moment may change only due to the motion of the charges constituting the dipole. But the motion of charges is accompanied by a current, and, hence, the variation of the polarization with time produces a current which is called the *polarization current.*

Let us consider the simplest case of a dipole consisting of two point charges (see equation 14.1)

$$\mathbf{p}_i = |e_i| \mathbf{l}_i \tag{55.9}$$

Let the radius vector of the positive charge be $\mathbf{r}_i^{(+)}$ and the radius vector of the negative charge be $\mathbf{r}_i^{(-)}$. Then we have (Fig. 68)

$$\mathbf{l}_i = \mathbf{r}_i^{(+)} - \mathbf{r}_i^{(-)} \tag{55.10}$$

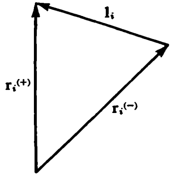


Fig. 68

and the expression (55.9) for the dipole may be put in the form

$$\mathbf{p}_i = |e_i|\mathbf{r}_i^{(+)} - |e_i|\mathbf{r}_i^{(-)} = e_i^{(+)}\mathbf{r}_i^{(+)} + e_i^{(-)}\mathbf{r}_i^{(-)} \tag{55.11}$$

where $e_i^{(+)}$ and $e_i^{(-)}$ are the positive and negative charges, respectively. Hence, equation (55.8) may be written

$$\mathbf{P} = \frac{1}{\Delta V} \sum_{\Delta V} (e_i^{(+)}\mathbf{r}_i^{(+)} + e_i^{(-)}\mathbf{r}_i^{(-)}) \tag{55.12}$$

When the charges move, $\mathbf{P}$ varies. Differentiating both sides of (55.12) with respect to time, we have

$$\dot{\mathbf{P}} = \frac{1}{\Delta V} \sum_{\Delta V} \left(e_i^{(+)} \frac{d\mathbf{r}_i^{(+)}}{dt} + e_i^{(-)} \frac{d\mathbf{r}_i^{(-)}}{dt} \right) = \frac{1}{\Delta V} \sum_{\Delta V} (e_i^{(+)}\mathbf{v}_i^{(+)} + e_i^{(-)}\mathbf{v}_i^{(-)}) \tag{55.13}$$

where $\mathbf{v}_i^{(+)}$ and $\mathbf{v}_i^{(-)}$ are the velocities of the positive and negative charge, respectively. The expression

$$\frac{1}{\Delta V}\sum_{\Delta V}(e_i^{(+)}\mathbf{v}_i^{(+)} + e_i^{(-)}\mathbf{v}_i^{(-)}) = \langle(\rho\mathbf{v})_{\text{pol}}\rangle \tag{55.14}$$

gives the mean density of the current associated with the change in the polarization. Thus

$$\dot{\mathbf{P}} = \langle(\rho\mathbf{v})_{\text{pol}}\rangle \tag{55.15}$$

which is the third term in the right-hand side of (55.5), describes the polarization current.

Magnetization Current. In neutral atoms, molecules and ions, molecular currents are produced by the motion of electrons in atoms. If there is no magnetization, then these molecular currents are distributed randomly, and their mean density, i.e., the mean value over a physically small volume, is zero. As a result of magnetization, however, the molecular currents become ordered, and their mean density is no longer equal to zero, but is defined by equation (20.15), i.e.

$$\langle\mathbf{j}_{\text{mol}}\rangle = \langle(\rho\mathbf{v})_{\text{mol}}\rangle = \operatorname{curl}\mathbf{I} \tag{55.16}$$

This is the second term on the right-hand side of equation (55.5). Thus, we have proved that the mean density of the microscopic current consists of the conduction, polarization and magnetization currents.

§56. Averaging the Charge Density

The mean charge density, clearly, is made up of the mean density of bound charges and the mean density of free charges

$$\langle\rho\rangle = \langle\rho_{\text{free}}\rangle + \langle\rho_{\text{bound}}\rangle \tag{56.1}$$

Averaging the density of the free charges gives the density denoted in (55.6) by ρ

$$\langle\rho_{\text{free}}\rangle = \rho \tag{56.2}$$

To find the mean density of the bound charges, we use equation (55.15), which describes the current produced by the motion of the bound charges. If the density of the bound charges in a given volume changes, then this must be the result of charges entering or leaving through the surface enclosing this volume. Hence, the law of conservation of the bound charges may be put in the form (compare the derivation of eq. (4.4))

$$\frac{\partial}{\partial t}\int_V \langle\rho_{\text{bound}}\rangle\, dV = -\oint_S \langle(\rho\mathbf{v})_{\text{pol}}\rangle\cdot d\mathbf{S} = -\oint_S \frac{\partial\mathbf{P}}{\partial t}\cdot d\mathbf{S} \tag{56.3}$$

where S is the surface enclosing the volume V. Since S is independent of time, the derivative with respect to time on the right-hand side of (56.3) may be taken outside the integral sign. Hence, we obtain

$$\frac{\partial}{\partial t}\int_V \langle\rho_{\text{bound}}\rangle\, dV = -\frac{\partial}{\partial t}\oint_S \mathbf{P}\cdot d\mathbf{S} \tag{56.4}$$

Integrating with respect to t, and using the fact that $\rho_{\text{bound}} = 0$ when $P = 0$, we find that

$$\int_V \langle\rho_{\text{bound}}\rangle\, dV = -\int_S \mathbf{P}\cdot d\mathbf{S} = -\int_S \operatorname{div}\mathbf{P}\, dV \tag{56.5}$$

by virtue of Gauss's theorem. In this equation, V is greater than or equal to a physically small volume, but is otherwise arbitrary. Since the volume of integration is arbitrary, it therefore follows from (56.5) that

$$\langle\rho_{\text{bound}}\rangle = -\operatorname{div}\mathbf{P} \tag{56.6}$$

Thus, using (56.2), equation (56.1) becomes

$$\langle\rho\rangle = \rho - \operatorname{div}\mathbf{P} \tag{56.7}$$

This relationship is the same as equation (55.6). Thus, it has been proved that, in fact, the mean density of the space charge is composed of the mean density of the free charges and the mean density of the bound charges.

We have, therefore, proved that Maxwell's equations are obtained by averaging the equations of the electron theory, and we have explained the physical meaning of the quantities occurring in phenomenological electrodynamics.

PART III

Theory of Relativity

ALL physical processes take place in space and time. Hence, physical theories are closely connected with the concepts of space and time.

The concepts of space and time have been historically developed from everyday experience. The fullest, and most physically meaningful formulation of these concepts was given by Newton, in the form of *absolute space* and *absolute time*. The theories of classical mechanics and gravitation, based on these concepts, agree exceptionally well with experimental data.

Everyday experience, on which the formulation of the concepts of absolute space and absolute time is based, relates to the practical activities of man. Man is concerned with the geometrical relationships between bodies, which can be observed with the naked eye. These bodies can move with velocities not very different from the velocities with which man himself can move about.

Man has not only formulated the concepts of space and time from everyday experience; he has also formulated, within the framework of these concepts, further physical concepts of mechanical motion as the change of position of a body in space, and of forces acting on a body, etc.

All these ideas and concepts have served very well as a

means of understanding objective reality as long as they have been used to describe phenomena which do not lie outside the framework within which those ideas and concepts have been formulated. But in the course of his practical experience, man has gone beyond this framework. The study of atomic phenomena has shown that many concepts and ideas, which have seemed so "natural" and "sure," became meaningless or simply erroneous when applied to atomic phenomena. The problem of a new approach to the understanding of atomic phenomena and corresponding new concepts and postulates has been solved by quantum mechanics. The study of motion at very high velocities has shown that the everyday concepts of space and time, apparently so "sure" and "natural," also needed fundamental revision. This revision has been provided by the theory of relativity.

Although man, from the very beginning, has made practical use of light, he did not study its velocity for a very long time. In antiquity, it had been deduced that the velocity of light traveling in a straight line was very great, by comparing it to the flight of an arrow, since the closer to a straight line the arrow's path becomes, the greater the speed of the arrow. But they had no real idea of the velocity of light, right up until the time of Galileo. Galileo himself tried to measure the velocity of light, using methods that are now obviously unsuitable. It is now known that the velocity of light is far greater than the limits of the speeds within which the concepts of absolute space and time were formulated. No wonder, then, that it was in studying the velocity of light within the framework of electrodynamics that scientists were first confronted with the insufficiency of the seemingly "sure" concepts of absolute space and time. The fundamental work by the founder of the theory of relativity, Albert Einstein, published in 1905, was called: "On the Electrodynamics of Moving Bodies."

The theory of relativity is not an electrodynamic theory; it is a general physical theory. However, there are good reasons why it is studied together with electrodynamics: first, its origins are closely connected with electrodynamics; second, electrodynamics is, historically, the first relativistically invariant theory; third, electrodynamics and the motion of

charged particles provide many examples of the general propositions of the theory of relativity.

The theory of relativity may be divided into two parts. In the first part, known as "the special theory of relativity," we study inertial coordinate systems in the absence of gravitational fields. The second part, which is called "the general theory of relativity," is devoted to the study of noninertial coordinate systems and gravitational fields. In the present volume, we shall consider only the special theory of relativity, which, for the sake of brevity, we shall call simply "the theory of relativity."

CHAPTER 12

Postulate of the Constancy of the Velocity of Light

§57. The Velocity of Light

Opinions of Ancient Philosophers. Ancient philosophers held two kinds of opinion about light. Plato (427 to 347 B.C.) put forward a theory of visual rays going out from the eye and, so to speak, "feeling" objects. Democritus (460 to 370 B.C.) was a supporter of the theory of emitted atoms impinging on the eye from the various objects. Aristotle (384 to 322 B.C.) also believed in the emission theory. However, the geometrical character given to optics by Euclid (300 B.C.), by establishing the rectilinear propagation of rays and the laws of reflection, made both points of view equally tenable. Later, the concept of emitted atoms became more prevalent, and it was assumed that the rays traveled at a very high velocity, or even instantaneously.

Opinions of the Founders of Classical Physics. Classical physics begins with Galileo Galilei (1564 to 1642). He assumed the velocity of light to be finite, but put forward no hypotheses about it. Descartes (1596 to 1650) believed that light was a pressure transmitted through a medium at infinite velocity. Thus, Descartes clearly believed that a medium was necessary for the transmission of light. Grimaldi (1618 to 1660) took a different view: according to him, light was a wave motion in a homogeneous medium. But the true founder of the wave theory of light was Christian Huygens (1629 to 1695), who described it to the Paris Academy of Sciences in 1678. Newton (1643 to 1727) was reluctant to commit himself concerning the nature of light "not wishing to fabricate a hypothesis," but he clearly held the cor-

puscular theory of emission, although he did not insist on its unqualified truth.

In 1675, Newton wrote: "I suppose light is neither aether, nor its vibrating motion, but something of a different kind propagated from lucid bodies. They, that will, may suppose it an aggregate of various peripatetic qualities. Others may suppose it multitudes of unimaginable small and swift corpuscles of various sizes, springing from shining bodies. . . ."†

Römer's Determination of the Velocity of Light (1676). The velocity of light was first measured by Römer in 1676. Observations of the occultations of the satellites of Jupiter showed that the apparent period of their revolution decreased when the earth was approaching Jupiter in the course of its annual motion, and increased when the earth was moving away from Jupiter. Römer realized that this effect was due to the fact that light travels with a finite velocity, and he calculated this velocity from his observations. Fig. 69 shows the position of the satellites of Jupiter immediately after an

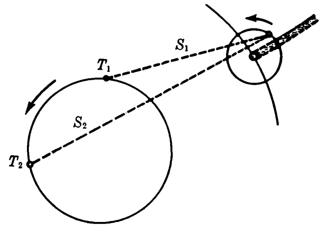


Fig. 69

occultation. Since the period of Jupiter's revolution about the sun is far greater than the period of the earth, Jupiter may be assumed stationary in this calculation. Let us suppose that at some instant t_1 a particular satellite is just emerging from Jupiter's shadow. To a terrestrial observer, it will appear to emerge at the instant

$$T_1 = t_1 + \frac{s_1}{c} \tag{57.1}$$

where s_1 is the distance between the earth and the point of emergence at the instant of observation, and c is the velocity of light. When, after one

† Quoted from Thomas Birch, *The History of the Royal Society of London*, pp. 254–255, as found in I. Bernard Cohen, *Isaac Newton's Papers and Letters on Natural Philosophy*, Harvard University Press, Cambridge, Mass., 1958, pp. 184–185.

revolution, the satellite once again emerges from the shadow, the terrestrial observer will observe it at the instant

$$T_2 = t_2 + \frac{s_2}{c} \tag{57.2}$$

Thus, according to the measurements of the terrestrial observer, the period of the satellite equals

$$T_{\text{obs}} = T_2 - T_1 = T_{\text{act}} + \frac{s_2 - s_1}{c} \tag{57.3}$$

where $T_{\text{act}} = t_2 - t_1$ is the true period of the satellite. Thus, due to a change in the distance of the earth from Jupiter $s_2 - s_1$, the observed period of a satellite of Jupiter differs from the true value. If a large number of observations are made, with the earth approaching and receding from Jupiter, the mean value of the observations will give the true period, since the averaged terms $(s_2 - s_1)/c$ may be positive or negative and they cancel out.

If we know T_{act}, we can use equation (57.3) to find the velocity of light

$$c = \frac{s_2 - s_1}{T_{\text{obs}} - T_{\text{act}}} \tag{57.4}$$

The values of s_2 and s_1 are known from astronomical calculations, since the motion of Jupiter and the motion of the earth are known. It is not difficult, therefore, to take into account the motion of Jupiter. The result obtained by Römer in this way for the velocity of light is close to the present day value.

Aberration of Light (Bradley, 1727). In the absence of wind, raindrops fall vertically, but the tracks they leave on the windows of a moving train are sloping. This is due to the addition of the vertical velocity of the drop and the horizontal velocity of the train. In the case of light, a similar phenomenon, called the *aberration of light*, is observed. Because of the aberration of light, the apparent direction of a star differs from the true direction (Fig. 70). The angle α between these directions is called the *angle of aberration*, and is related to the velocity of light c and the perpendicular component of the velocity $v_\perp$ by the equation

$$\tan \alpha = \frac{v_\perp}{c} \tag{57.5}$$

In practice, the aberration of light is observed as follows. In every observation throughout the year, the axis of a telescope is oriented in an identical manner in space with respect to the stellar sky, and the position of the star's image in the focal plane of the telescope is noted. In the course of the year this image describes an ellipse. If we know the dimensions of this

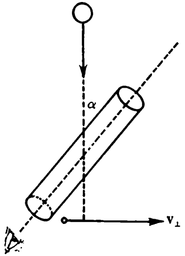


Fig. 70

ellipse and other observational data, we can determine the angle of aberration. If we know $v_\perp$, and measure α, then we can determine the velocity of light. The results of this method confirmed the value obtained by Römer.

Different Interpretations of the Velocity of Light. The value of the velocity of light was established. The question then arose, however, on what does this velocity depend? The answer to this question depended on the view taken concerning the nature of light.

If light is a wavelike motion of a homogenous medium, then its velocity with respect to this medium should be constant, governed by the properties of the medium. The velocity of light with respect to its source and to an observer should, however, depend on the velocity of the source and the observer with respect to the medium.

If light represents the flow of swift corpuscles emitted from a source, then it would be natural to suppose that these corpuscles are subject to the laws of nature, in particular to the law of inertia. Hence, the velocity of these particles with respect to an observer should be the resultant, by the parallelogram rule, of the velocity of the corpuscles with respect to the source and the velocity of the source with respect to the observer.

Hypothesis of an All-Pervading Ether and Absolute Velocity. The corpuscular theory had the backing of the authority of Newton, so that Huygens' wave theory, although it had its supporters, was ignored for more than a century. However, at the beginning of the nineteenth century, new discoveries in optics changed the position radically. In 1801, Young established the principle of interference, and used it to explain the colors of thin plates. However, the ideas put forward by Young were of a qualitative nature, and were not accepted widely. The coup-de-grace was given

to the corpuscular theory in 1818, when Fresnel used the wave theory to solve the problem of diffraction. All attempts to solve this problem in terms of the corpuscular theory were unsuccessful. The basis of Fresnel's ideas was the combination of Huygens' principle of elementary waves and Young's principle of interference. During the next few years, the corpuscular theory was completely eliminated from science, and the concept of light as a wave motion in a medium became generally accepted. The medium, which was assumed to pervade the whole universe was given the name of *ether*. The problem was now reduced to constructing a theory of light in the form of oscillations of the ether. Later, the role of the ether was extended to include other phenomena (gravitation, magnetism, electricity). Many notable scientists of the last century worked on the theory of the ether, but their work is now only of historical interest. Nowadays, we only refer to the ether to clarify the concept of *absolute velocity* and the methods of finding it.

The ether was assumed to fill the entire universe. The ether itself was assumed to be stationary. The velocity of light with respect to the ether was taken to be constant, defined by the properties of the ether. Bodies moved with respect to the stationary ether which filled all space. Thus, the motion of bodies with respect to the ether would be absolute in nature, and different from the motion of bodies with respect to each other. In general, if a body A moves with respect to a body B with a velocity $\mathbf{v}$, we can change the value of this relative velocity by applying a force to B. But on the ether theory, we could only change the motion of A with respect to the ether by applying a force to A. Therefore, the velocity with respect to the ether was called the absolute velocity. This absolute velocity for a given body would be independent of the motion of all other bodies. It would have meaning even in the absence of all other bodies. The question now was, how to measure it.

Measurement of the "Absolute Velocity." Since the velocity of light with respect to the ether was assumed to be constant, this velocity should vary relative to bodies moving in the ether, and should depend on the velocity of these bodies with respect to the ether. Hence, by measuring the velocity of bodies with respect to light, or, which amounts to the same thing, the velocity of light with respect to moving bodies, we should be able to determine the velocity of a body with respect to the ether (the velocity of light with respect to the ether was assumed to be known). This situation is completely analogous to that of an oarsman in a boat; if he measures the speed of the boat with respect to the waves and knows the speed of the waves with respect to still water, he can find the velocity of the boat with respect to the water.

Such an attempt to determine the absolute velocity of the earth was made by Michelson in 1881 and 1887.

§58. Michelson's Experiment

Principle and Method of the Experiment. The basis of Michelson's experiment is to compare the velocity of light along two paths, one of which lies in the direction of motion of the earth through the ether, and the other perpendicular to that motion. The apparatus used by Michelson is shown in Fig. 71.

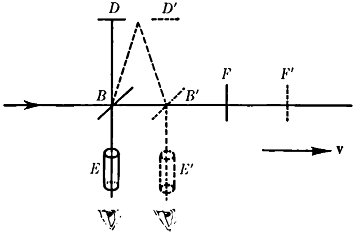


Fig. 71

A monochromatic beam from a source A strikes a semitransparent plate B, inclined at an angle of 45°. Here, the beam is split into two coherent beams. One beam is reflected from the plate to a mirror D, while the other passes through the plate to a mirror F. The mirrors D and F reflect the beams back to B. A part of the beam from D passes through B and reaches an interferometer E together with a part of the beam reflected from F and B.

Calculation of the Path Difference Between the Beams. Let us assume that the apparatus is moving in the direction of the arm BF $= l_1$ with a velocity v. Let us denote the velocity of light with respect to the ether by c. When the beam travels from B to F, it is moving in the direction of motion of the apparatus. Hence, the velocity of light with respect to the apparatus is $c - v$, and the time taken for the light to travel along BF equals

$$t_{BF}^{(1)} = \frac{l_1}{c - v} \tag{58.1}$$

The time taken for the beam to travel back along this path after reflection from F is

$$t_{\mathrm{FB}}^{(1)} = \frac{l_1}{c+v} \tag{58.2}$$

since now the light travels in the opposite direction to the apparatus, and the velocities are therefore added.

Thus, the total time taken for the light to travel from B to F and back again is

$$t_{\parallel}^{(1)} = t_{\mathrm{BF}}^{(1)} + t_{\mathrm{FB}}^{(1)} = \frac{2l_1}{c}\frac{1}{1-\dfrac{v^2}{c^2}} \tag{58.3}$$

The superscript 1 denotes the time taken by the light to travel along various paths when the apparatus has the stated orientation with respect to its direction of motion. The superscript 2 denotes time intervals for the other orientation, when the direction of motion of the apparatus is along BD.

To determine the time taken for the light to travel along BDB we shall consider the beam reflected from B and traveling to D. The velocity of light must be resolved into two components: a component in the direction of motion of the apparatus, equal to v, and a perpendicular component $c_\perp$ from B to D. Hence, we may write

$$c^2 = c_\perp^2 + v^2 \tag{58.4}$$

The time taken for the beam to travel along BD $= l_2$ equals

$$t_{BD}^{(1)} = \frac{l_2}{c_\perp} = \frac{l_2}{\sqrt{c^2 - v^2}} = \frac{l_2}{c}\frac{1}{\sqrt{1-\dfrac{v^2}{c^2}}} \tag{58.5}$$

The velocity of the beam in the opposite direction is also $c_\perp$, and, hence, the time taken to travel along DB is the same. Hence, the total time for the beam to travel from B to D and back again is

$$t_\perp^{(1)} = \frac{2l_2}{c}\frac{1}{\sqrt{1-\dfrac{v^2}{c^2}}} \tag{58.6}$$

The speed of the earth along its orbit round the sun is 30 km/sec. Hence, if the apparatus is stationary on the earth, the ratio $(v/c)^2$ is of the order

$$\left(\frac{v}{c}\right)^2 \approx 10^{-8} \tag{58.7}$$

Since $(v/c)^2$ is a small quantity, we can expand (58.3) and (58.6) in series in terms of $(v/c)^2$ and take only the first term. We thus obtain

$$t_{\parallel}{}^{(1)} \approx \frac{2l_1}{c}\left(1 + \frac{v^2}{c^2}\right) \qquad t_{\perp}{}^{(1)} \approx \frac{2l_2}{c}\left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) \tag{58.8}$$

Hence, the difference between the travel times is equal to

$$\Delta t^{(1)} = t_{\parallel}{}^{(1)} - t_{\perp}{}^{(1)} = \frac{2}{c}\frac{v^2}{c^2}\left(l_1 - \frac{l_2}{2}\right) + \frac{2}{c}(l_1 - l_2) \tag{58.9}$$

We now turn the apparatus through 90° so that the direction of motion of the apparatus is now along BD, and BF is perpendicular to the direction of motion. The time difference for the two paths is determined in exactly the same way, but now instead of l_1, we have l_2, and vice versa. Hence, instead of (58.9) we have

$$\Delta t^{(2)} = t_{\parallel}{}^{(2)} - t_{\perp}{}^{(2)} = \frac{2}{c}\frac{v^2}{c^2}\left(l_2 - \frac{l_1}{2}\right) + \frac{2}{c}(l_2 - l_1)$$

Thus, the total time difference when the apparatus is turned through 90° is

$$\Delta t = \Delta t^{(1)} + \Delta t^{(2)} = \frac{l_1 + l_2}{c}\frac{v^2}{c^2} \tag{58.10}$$

Result of Michelson's Experiment. The true motion of the apparatus with respect to the (postulated) ether is unknown. We cannot, therefore, actually orient the apparatus so that one arm of it lies along the direction of motion. Therefore, the apparatus was slowly rotated during the experiment. Whatever the direction of motion of the apparatus with respect to the ether, in every 360° turn each arm would be oriented twice along the direction of motion, and twice perpendicular to it. The interference bands were observed with an interferometer. If the difference between the beam paths changed during the process of rotation, the interference bands would change their position in the field of view. The change in the difference between the beam paths could then be calculated from the displacement of the interference bands, and the velocity of the apparatus with respect to the ether be determined.

This experiment was carried out by Michelson in 1881, and later, in 1887, by Michelson and Morley. To increase the magnitude of the effect, Michelson and Morley, in the later experiment, used multiple reflection of a beam from mirrors, and, thus, increased the effective lengths l_1 and l_2 to about 10 m. The wavelengths of visible light lie within the range (0.4 to 0.75) $\times$ 10^{-4} cm. The lag given by equation (58.10), and expressed in the form of a wavelength shift, is

$$\Delta\lambda = \Delta t c = (l_1 + l_2)\frac{v^2}{c^2} \approx (l_1 + l_2)10^{-8} \tag{58.11}$$

where $(v/c)^2$ is taken from (58.7), using the speed of rotation of the earth

around the sun. Therefore, for the wavelength $\lambda = 0.5 \times 10^{-4}$ cm, the relative shift of the interference bands should be

$$\frac{\Delta\lambda}{\lambda} = \frac{(l_1 + l_2)10^{-8}}{\lambda} = \frac{l_1 + l_2}{5 \times 10^3} \tag{58.12}$$

In the tests of 1887, the effective length was 11 m, and the expected displacement of the interference lines should have been easily observed. The experiment was so designed that a 3 km/sec velocity of the apparatus with respect to the ether would have been noticed. However, no effect was observed at all. The test was carried out again, with even greater precision, in 1905, but again gave a negative result. This implies that the earth does not move with respect to the postulated ether.

Interpretation of the Result of Michelson's Experiment Using the Ether Hypothesis. To account for the result of Michelson's experiment, two suggestions were made: (1) when close to a massive body, such as the earth, the ether moves together with the body, i.e., it is dragged by the motion of the body, so that close to such bodies no "ether wind" can be observed; (2) the dimensions of bodies moving in the ether change in such a way that the expected difference (58.10) cannot be observed.

The hypothesis of the ether drag had to be discarded because it contradicted other observations, in particular, it did not agree with the phenomenon of the aberration of light. The second assumption, put forward by Lorentz and Fitzgerald, successfully explained the absence of the lag. Comparison of (58.3) and (58.6) shows that when $l_1 = l_2 = l$ the times taken to travel along the paths collinear with, and perpendicular to the direction of motion are equal if the length of the arm along the direction of motion contracts to become

$$l' = l\sqrt{1 - \frac{v^2}{c^2}} \tag{58.13}$$

Thus, it is possible to explain Michelson's result by assuming a contraction, given by (58.13), for all bodies in the direction of their motion.

This explanation is, however, unsatisfactory, since it results in an illogical conclusion regarding the velocity of light. It is assumed that the velocity of light is constant with respect to the ether, and variable with respect to bodies moving in the ether; however, measurement of the velocity of light with respect to various moving bodies always gives the same result. In short, the velocity of light with respect to bodies is variable, but the results of measurements show that it is constant. It is clear that in this situation the hypothesis that the velocity of light is variable is quite untenable. Hence, we must assume that the velocity of light is constant, and then Michelson's result is explained quite naturally.

§59. The Ballistic Hypothesis

There is another method of explaining Michelson's result: we can leave out the ether altogether and assume that light represents the flow of material corpuscles, i.e., we can return to Newton's original point of view. We naturally assume that these "corpuscles" obey the law of inertia. Consequently, the velocity of light is added to the velocity of the source by the parallelogram rule, but the velocity of light relative to the source is constant.

Since the velocity of light is the same in all directions with respect to the source, according to the ballistic hypothesis it is impossible to expect a positive result from Michelson's experiment. Hence, the ballistic hypothesis explains Michelson's result in a completely natural manner. However, this hypothesis proves to be inconsistent, as described below.

Inconsistency of the Ballistic Hypothesis. The inconsistency of the ballistic hypothesis follows from the astronomical observations of double stars, first discussed by de Sitter in 1913.

A double star is a pair of stars, comparatively close together, moving about their common center of gravity. If one of the stars has a considerably greater mass than the other, we may assume that the star of greater mass is at rest, and that the other revolves around it. Such double stars occur in fairly large numbers. The speed of the stars may be measured, using the Doppler effect, and the elements of the orbit calculated accordingly. This shows that the component stars obey Kepler's laws of motion. No peculiarities are observed in the motion of double stars. However, if the ballistic hypothesis were correct, the motion of double stars would appear to be very peculiar.

Let us assume that we observe a double star from a sufficiently great distance s. For simplicity, we shall assume that the less massive star moves in a circle with a speed v about the other star, which may be considered to be stationary (Fig. 72). Let us consider a ray of light emitted at the instant when the star is at B, moving away from the observer. This ray

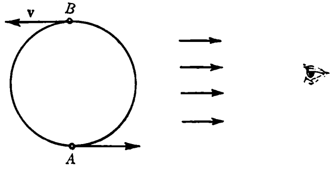


Fig. 72

will move towards the observer with velocity $c - v$. If it was emitted at time t_1 then it will reach the observer's eye at time

$$T_1 = t_1 + \frac{s}{c - v} \tag{59.1}$$

After half the period of revolution, $T/2$, the star emits a ray from A, while moving towards the observer. This ray moves towards the observer with velocity $c + v$, and reaches him, therefore, at time

$$T_2 = t_1 + \frac{T}{2} + \frac{s}{c + v} \tag{59.2}$$

If the distance is sufficiently large, the ray of higher velocity will overtake the ray emitted by B. This will occur at a distance s for which $T_2 = T_1$, and which may easily be found from (59.1) and (59.2). At greater distances the ray from A may overtake the ray from B emitted during the previous revolution of the star, with the result that an observer sufficiently far away would see the star at several points of the orbit simultaneously.

Thus, if the ballistic hypothesis were correct, astronomers observing double stars would see a rather complicated picture. In fact, they do not observe anything of the kind. What they do see agrees with the assumption that double stars obey Kepler's laws of motion, and that the velocity of light is constant and is not added to the velocity of the source, as the ballistic hypothesis would require. Thus, the observations of double stars disprove the ballistic hypothesis.

We therefore deduce that the velocity of light is independent of the velocity of the source. The result of Michelson's experiment shows that it is likewise independent of the velocity of the observer. Thus, all the observations we have discussed show that the velocity of light is constant, and independent of the velocities of both source and observer.

§60. Fizeau's Experiment

Incompatibility of c = const with the Usual Concepts. The concept that the velocity of light is constant is totally incompatible with the concepts of classical mechanics and everyday experience. Let us consider a train moving at 100 km/hour. If a man walks along the corridor toward the front of the train at 5 km/hr (with respect to the train), then he is moving at 105 km/hr with respect to the rails. This result is in complete agreement with the usual concepts of space and time, which, in this case, are expressed by the classical formula for the addition of velocities.

Let us now consider a rocket moving with a velocity of 100,000 km/sec with respect to the earth. If a ray of light is moving with velocity 300,000

km/sec with respect to the rocket, and in the same direction as the rocket, then we would expect, from the usual assumptions, the velocity of the ray with respect to the earth to be 400,000 km/sec. But this is not so; from the constancy of the velocity of light, it follows that the velocity of this ray with respect to the earth is also 300,000 km/sec. Thus, the ray of light moves with respect to the earth and with respect to the rocket at the same velocity of 300,000 km/sec, although the rocket is moving with respect to the earth with the velocity of 100,000 km/sec. This is completely inexplicable in the usual concepts of space and time. To explain it, we must change these concepts in such a way that on the basis of our new concepts, these results are quite natural. Such a change in the concepts of space and time is made in the theory of relativity.

Principle of Fizeau's Experiment. Long before the establishment of the concept of the constancy of the velocity of light, a physical experiment was known which demonstrated a strange law for the addition of large velocities, comparable with the velocity of light. This was Fizeau's experiment, carried out in 1860. The result of the experiment did not seem remarkable at the time, since it was explained by the hypothesis that the ether is partially carried along with the motion. The true meaning of the result was only understood much later.

The principle of Fizeau's test is the measurement of the velocity of light in a moving medium. Let $u' = c/n$ be the velocity of light in a medium whose refractive index is n. If the medium is itself moving with a velocity v, then, with respect to an observer at rest, the velocity of light should be $u' \pm v$, depending on whether the light is moving in the same or the opposite direction as the medium. In this experiment, Fizeau compared the velocities of rays of light moving in the same direction as the medium and opposite to it.

The layout of the experiment is shown in Fig. 73. A monochromatic ray from a source A strikes a semitransparent plate B, and is split into two coherent rays. The ray reflected from B moves along the path BKDEB (K, D, and E are mirrors) while the ray transmitted by B moves along the path BEDKB, i.e., in the opposite direction to the reflected ray. The first ray, having returned to B, is partially reflected by B, and reaches an interferometer F. The second ray is partially transmitted on returning to B, and also reaches F. Both rays move along the same path, and the sections BE and KD of this path pass through a liquid. If the liquid is at rest, then the paths of both rays are completely equivalent; it takes the same amount of time for both rays to travel over the whole path, and there is no difference between them.

However, when the liquid is moving, the paths of the rays are not

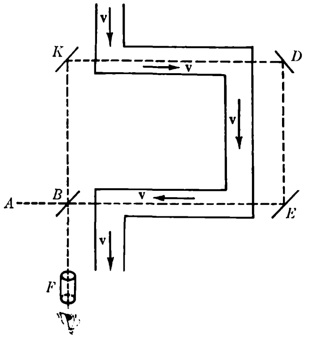


Fig. 73

equivalent; one moves with the flow of the liquid, and one moves against the flow. The difference may be determined from the interference pattern, and, hence, the velocity of light through the liquid may be calculated, since the velocity of light along the other parts of the path and all the relevant lengths are known.

Calculation of the Path Difference Between the Rays. We shall use l to denote the total length of the sections of the path of light in the liquid. The time taken for the light to travel the whole path, excluding the parts in the liquid, will be denoted by t_0. The velocity of the ray of light moving with the flow of the liquid will be denoted by $u^{(+)}$, and that moving against the flow of the liquid by $u^{(-)}$. We may write these velocities in the form

$$\begin{aligned} u^{(+)} &= u' + kv \\ u^{(-)} &= u' - kv \end{aligned} \tag{60.1}$$

where k is an undefined coefficient to be determined in the experiment. If $k = 1$, we have the classical formula for the addition of velocities. If $k \neq 1$, there is a deviation from the classical formula.

The times taken for the first and second rays to travel the entire path are, respectively

$$\begin{aligned} t_1 &= t_0 + \frac{l}{u' + kv} \\ t_2 &= t_0 + \frac{l}{u' - kv} \end{aligned} \tag{60.2}$$

Hence, we obtain the following expression for the time difference between the rays

$$\Delta t = t_2 - t_1 = l \frac{2kv}{u'^2 - k^2v^2} \tag{60.3}$$

Having found this difference from the displacement of the interference bands, and knowing l, v, u', we can evaluate k from the above equation.

Result of the Experiment. Fizeau's experiment gives the value of k as

$$k = 1 - \frac{1}{n^2} \tag{60.4}$$

where n is the refractive index of the liquid. Thus, the velocity of light in the liquid and the velocity of the liquid are not added according to the classical formula. From the point of view of the usual concepts, this result is as remarkable as the assertion that the velocity of light *in vacuo* is constant. However, at the time of Fizeau's experiment there seemed nothing remarkable in this result. Long before Fizeau's experiment, Fresnel had shown that a body moving through the ether partially carried the ether along with it, and this exactly corresponded to the result obtained later by Fizeau's experiment.

It was only after the formulation of the theory of relativity that it became clear that Fizeau's experiment was the first experimental proof of the inaccuracy of the classical law of the addition of velocities.

§61. Postulate that the Velocity of Light Is Constant

Why c = const Is a Postulate. All the experimental facts cited in the preceding paragraph point to the conclusion that the velocity of light is constant. This is incompatible with the Newtonian concepts of space and time. Hence, the postulate that the velocity of light is constant entails a considerable change in our concepts of space and time.

However, every experimental fact is known with only a certain degree of precision, and, hence, the result c = const is deduced from the experimental data with only a certain degree of precision. Strictly speaking, therefore, the assertion that the velocity of light is constant exceeds the bounds of experimental precision, i.e., it is a postulate. Therefore, here and later we shall regard as a postulate the assertion that the velocity of light is constant.

Absence of Direct Experimental Confirmation. At present there is no direct experimental confirmation of the postulate c = const. Experiments in which the light passes through a transparent medium or is reflected before measurement cannot fill this role, as the refraction or reflection

of light has a very considerable effect on its subsequent motion. Hence, experiments of this type cannot be considered as providing direct experimental proof that the velocity of light is constant. Nor can the observations of double stars be considered as direct experimental proof. It could be assumed that immediately after emission, the velocity of the light is added to the velocity of the star. Then, during its propagation in space, the action of the interstellar medium, gravitational fields, etc., may gradually make the velocity of light a constant value, characteristic of the conditions in the interstellar space. If the time for "equalization of velocities" is fairly short, then any dependence of the velocity of light on the velocity of the source could not be found by observations on double stars.

All the experimental facts and observations which we have at present are indirect proofs that c = const. However, modern experimental techniques are such as to make direct experimental proof possible, and such proofs may well be obtained in the not too distant future. These direct proofs would be of the greatest importance, since it is always desirable to have at least one direct experimental proof of any assertion in physics.

Justification of the Reliability and Correctness of the Postulate. Nowadays, there are no doubts as to the correctness of the postulate. The entire special theory of relativity is founded on the postulate that the velocity of light is constant and on the principle of relativity. Hence, every confirmation of the deductions from the theory of relativity carries with it a confirmation of the correctness of the initial propositions. There are a great many experimental confirmations of the special theory of relativity, and the deductions from the theory of relativity are used in practical engineering. There are no doubts, nowadays, of the correctness of the special theory of relativity, hence, there are no doubts either of the postulate that the speed of light is constant.

CHAPTER 13

The Principle of Relativity

§62. Frames of Reference

Coordinate Systems. In physics, it is always necessary to specify, first of all, where and when a given phenomenon occurs. For example, when we study the mechanical motion of a material point, we must specify its locus at various instants of time. An electromagnetic field is defined by giving values of **E** and **H** at various points in space at various instants of time. Thus, the definition of the space coordinates and of time is the first step in the description and study of physical phenomena and processes.

An empty space devoid of matter is an abstraction, and it is impossible to give a method for distinguishing one part of the empty space from another. We may say that a continuous empty space does not contain its own measure. The question of measuring space becomes physically meaningful only when there is matter present, in which case the points in space differ from one another because of their different positions with respect to matter. Thus, matter in the form of a body or bodies may serve as a reference system. A coordinate system may be associated with a reference system, e.g., a rectangular Cartesian system. By choosing a unit of length, we may define uniquely the coordinates of any point by measuring the position of the point with respect to the chosen coordinate system. Once a coordinate system and a standard of measurement are introduced, the problem of determining the position of a point no longer belongs to physics, but to geometry. Every point in space is described by three numbers (e.g., x, y, z in the Cartesian system), and conversely, any chosen set of three numbers describes a point in space.

Transformation of Coordinates. A coordinate system may be chosen in a countless number of ways. It may be related, for example, to the earth or to the sun, or even to the stellar system, but in all these systems, points in space will be described by a set of numbers proper to them. If we know the coordinates of a point in one system, we can find its coordinates in another system, using the appropriate transformation rule. Thus, if we have two rectangular coordinate systems, such that one is turned through an angle α with respect to the other (Fig. 74), then, as we know from geometry, the transformation of the coordinates will take the form

$$x' = \cos\alpha\, x + \sin\alpha\, y$$
$$y' = -\sin\alpha\, x + \cos\alpha\, y \qquad (62.1)$$

Using such a transformation, we may pass from one coordinate system to another.

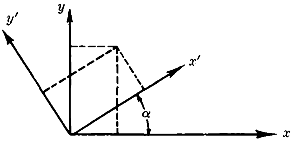


Fig. 74

Measurement of Time. Coordinate systems may move with respect to one another. Thus, using the example of Fig. 74, the x', y' coordinates may rotate with respect to the x, y, coordinates with an angular velocity ω. Then α depends on time: $\alpha = \omega t + \alpha_0$, where $\alpha = \alpha_0$ at $t = 0$. Hence, the transformation (62.1) will in this case be time-dependent. In speaking of the coordinates of a point in one of the coordinate systems of Fig. 74, we must specify the instant to which the coordinates refer. If we do not indicate the time, then we will not know how to find the point in space corresponding to given coordinates. Since in physics all coordinate systems are, generally speaking, moving with respect to one another, the indication of the time of an event is as necessary as the indication of the point at which it occurs.

Time is measured by means of some periodic process. Consideration of all periodic processes enables us to establish standards of time. We shall speak of "clocks" by means of which time is measured, meaning the standard has been chosen for the measurement of time.

Let us take a coordinate system in which the position of every point in space is described by the coordinates of that point. To describe the time

of some event occurring at some point, we must imagine a clock at that point, i.e., we must suppose that there is a clock at every point of the system under consideration, so that the time of an event at a point may be determined from the clock. Thus, every point of a given coordinate system is described by the space coordinates and the time shown by the clock at that point. It must be stressed once more that we only imagine a clock to be placed at a point in order to give a clear physical meaning to the statement that this or that event takes place at such and such a point and at such and such a time.

Clocks measure time, which passes irrespective of whether or not there is a clock at a given point. However, to work with time as a physical quantity we must be able to measure it. This measurement of time is carried out by means of a clock.

Synchronization of Clocks. One point has not yet been made clear in the question of measuring time: how to synchronize clocks at different points, i.e., how to compare the origin of the frame of reference for clocks at different points. This may be accomplished by means of light signals using the postulate $c =$ const. Let us suppose that at time t a spherical light wave is radiated from the origin of coordinates. At the instant when the light reaches a point at a distance r from the origin, the clock at this point should show time $t_r = t + r/c$. If this is so, then the clock at this point is synchronized with the clock at the origin. It is easy to see that if a clock A is synchronized with a clock B, and the clock B with a clock D, then clock A is synchronized with clock D.

Meaning of the Transformation of Time. To enable us to compare the time in different frames of reference, we need a rule for transforming not only the coordinates but also the time. As an example, let us consider the transformation (62.1) again. Let us assume that the coordinate systems are at rest with respect to each other (Fig. 74), i.e., $\alpha =$ const. However, the origins of time in the frames of reference $S = S(x, y, z, t)$ and $S' = S'(x', y', z', t')$ are, generally speaking, different. Suppose, for example, that a clock in S′ is fast with respect to a clock in S by an amount t_0. Then the time transformation takes the form

$$t' = t + t_0 \tag{62.2}$$

The meaning of equation (62.2) is as follows. At every point in space, there is a clock belonging to S′ and a clock belonging to S. The S′ clock shows time t' at the instant when the S clock *at the same point* shows time t. In our example, the difference in the readings of the clocks is purely formal. We can change the time origin in one of the frames of reference, and make the readings of the two clocks identical

$$t' = t \tag{62.3}$$

The transformation (62.2) is written down only to make the meaning of the time transformation clearer: we are comparing the readings of the clocks belonging to different coordinate systems which are at the same point at the instant of comparison.

Inertial Coordinate Systems. Among all the imaginable coordinate systems, there is a class of systems distinguished by its special simplicity—the *inertial coordinate systems.*

According to Newton's first law of motion, a body, sufficiently removed from other bodies and not experiencing an external force, is in a state of rest, or of uniform rectilinear motion. This raises the question of the coordinate system in which the motion is measured. It is clear that if we consider the motion of a sufficiently distant body, subject to no external force with respect to, e.g., a rotating coordinate system, then in this coordinate system the motion will be neither rectilinear nor uniform. Newton's law of inertia, formulated above, is correct in coordinate systems associated with bodies not subjected to the action of external forces. Coordinate systems in which Newton's law of inertia holds are called inertial coordinate systems. All these coordinate systems are in uniform rectilinear translatory motion with respect to one another.

In Newtonian mechanics, inertial systems occupy a privileged place among all the imaginable coordinate systems. This is due to the fact that Newton's laws hold only for inertial coordinate systems. The special theory of relativity, which is discussed in this book, also deals *only* with inertial coordinate systems. Hence, in the following discussion, a "coordinate system" will always mean an "inertial coordinate system." As a rule, we shall use rectangular Cartesian systems.

Invariance and Covariance. The transition from one frame of reference of coordinates and time to another is accomplished by means of transformation formulas. In such a transformation, the various physical quantities involved will, generally speaking, change in numerical value, and the equations describing the various physical processes will, generally speaking, also change in form. But this is not always so.

If the value of some quantity remains the same after transition from one frame of reference to another, then we say that the quantity is *invariant* under this transformation. Such quantities are also called the invariants of a given class of transformations.

If the form of some equation remains the same after transition from one frame of reference to another, i.e., if the equation in the variables of the new system has the same form as in the variables of the old system, then

we say that the equation is *covariant* under this transformation. If not only the form, but also the numerical values of the individual terms of the equation remain unchanged, i.e., if the terms of the equation are invariant, then we say that the equation is invariant under this transformation.

§63. The Principle of Relativity in Classical Mechanics

Galileo's Transformation. If coordinate systems are at rest with respect to one another, then the question of transforming the coordinates is purely geometrical. It may be reduced to a rotation of the coordinate systems with respect to one another, and described by equations of the form of (62.1) and a translation of the origin. The equation for the time transformation has the form (62.3), since the transformation under discussion is purely geometrical and is not related to any time-dependent quantities.

The transformation of coordinates becomes a physics problem when the coordinate systems are moving with respect to one another. Let us consider inertial rectangular Cartesian systems. Such systems move in uniform rectilinear motion with respect to one another. All motion is relative, hence, any of the systems may be considered to be at rest. Quantities measured with respect to a coordinate system at rest will be denoted by letters without a prime, and those measured with respect to a moving coordinate system will be denoted by the same letters, but with primes.

Let us consider the expressions for the transformation of coordinates when the x axis of the moving system (S′) coincides with the x axis of the system at rest (S), and the y' and z' axes are parallel to the y and z axes, respectively. The moving system S′ moves with a velocity v in the positive x direction (Fig. 75). This special choice of the orientation of the axes of the systems does not restrict the generality of the transformation, since

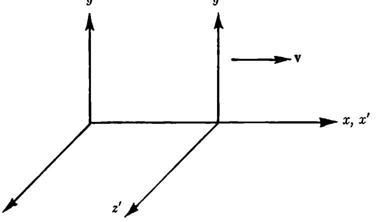


Fig. 75

given any mutual orientation of the sets of axes we can always, by the purely geometrical process of rotating one set of axes and displacing its origin, obtain the orientation just described. Since the necessary geometrical transformation is known, it is sufficient to consider the case shown in Fig. 75.

Let us suppose that at an instant $t = 0$ the origins coincide. Then the transformation formulas may be written

$$x' = x - vt \qquad z' = z$$
$$y' = y \qquad t' = t \tag{63.1}$$

These expressions are called the *Galilean transformation.*

Clearly, we may take S′ as the system at rest, and S as the moving system. In this case, we must interchange the quantities with and without a prime in equation (63.1) and change the sign of the velocity, since S is moving with respect to S′ in the negative x' direction. Hence, taking S′ as the system at rest, we may rewrite the Galilean transformation

$$x = x' + vt' \qquad z = z'$$
$$y = y' \qquad t = t' \tag{63.2}$$

Invariance of Length. Let a rod be placed in the S′ system so that the coordinates of its ends are x'_1, y'_1, z'_1 and x'_2, y'_2, z'_2. This means that the length of the rod in S′ is equal to

$$l' = \sqrt{(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2}$$

In the S system, the rod moves with a velocity v, and the coordinates of the ends are time-dependent. The *length* of a moving rod is the distance between the coordinates of its ends at some given instant. Thus, to measure a moving rod, we must note the position of its ends simultaneously, i.e., at the same instant of time. Then the distance between the noted positions will be the length of the moving rod. Let the positions of the ends be noted at a time t_0. Using (63.1), the coordinates of the ends will be

$$\left.\begin{aligned} x'_1 &= x_1 - vt_0 \qquad x'_2 = x_2 - vt_0 \\ y'_1 &= y_1 \qquad y'_2 = y_2 \\ z'_1 &= z_1 \qquad z = z_2 \end{aligned}\right\} \tag{63.3}$$

Hence, it follows that

$$x'_2 - x'_1 = x_2 - x_1 \qquad y'_2 - y'_1 = y_2 - y_1 \quad z'_2 - z'_1 = z_2 - z_1$$

Thus, we have

$$\begin{aligned} l &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \\ &= \sqrt{(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2} = l' \end{aligned} \tag{63.4}$$

i.e., the length of the rod is the same in both systems, i.e., length is invariant under the Galilean transformation.

Invariance of Time Interval. The invariance of an interval of time is proved using the expression $t = t'$. Suppose that at some point of the moving system events occur at t'_1 and t'_2. The time interval between them is

$$\Delta t' = t'_2 - t'_1 \tag{63.5}$$

In S, these events, according to equation (63.2), occur at the instants

$$t_1 = t'_1 \qquad t_2 = t'_2 \tag{63.6}$$

and, consequently, the interval between them is

$$\Delta t = t_2 - t_1 = t'_2 - t'_1 = \Delta t' \tag{63.7}$$

i.e., time intervals are invariant under the Galilean transformation.

Addition of Velocities. Let us suppose that a particle moves in S′ so that its coordinates are defined with respect to time by

$$x' = x'(t') \qquad y' = y'(t') \qquad z = z'(t') \tag{63.8}$$

The components of velocity of this point in S′ are

$$u'_x = \frac{dx'}{dt'} \qquad u'_y = \frac{dy'}{dt'} \qquad u'_z = \frac{dx'}{dt'} \tag{63.9}$$

In S′, using the Galilean transformation (63.2) with $v = \text{const}$, the law of change of the coordinates with respect to time becomes

$$\begin{aligned} &x(t) = x'(t') + vt' \qquad z(t) = z'(t') \\ &y(t) = y'(t') \qquad t = t' \end{aligned} \tag{63.10}$$

Hence, the components of velocity of the point in S are given by

$$\left.\begin{aligned} u_x &= \frac{dx}{dt} = \frac{dx'}{dt} + v\frac{dt'}{dt} = \frac{dx'}{dt'} + v\frac{dt'}{dt'} = u'_x + v \\ u_y &= \frac{dy}{dt} = \frac{dy'}{dt} = \frac{dy'}{dt'} = u'_y \\ u_z &= \frac{dz}{dt} = \frac{dz'}{dt} = \frac{dz'}{dt'} = u'_z \end{aligned}\right\} \tag{63.11}$$

which are the classical formulas for the addition of velocities.

Invariance of Acceleration. Differentiating equations (63.11) with respect to time, and remembering that $dt = dt'$, we obtain

$$\left.\begin{aligned} \frac{d^2x}{dt^2} &= \frac{d^2x'}{dt'^2} \\ \frac{d^2y}{dt^2} &= \frac{d^2y'}{dt'^2} \\ \frac{d^2z}{dt^2} &= \frac{d^2z'}{dt'^2} \end{aligned}\right\} \tag{63.12}$$

These equations show that acceleration is invariant under the Galilean transformation.

Invariance of Newton's Equations. In Newtonian mechanics, forces are the result of the interaction of material bodies. It is assumed that these forces depend only on the distances between bodies and are independent of velocity and acceleration. Such forces occur in electrostatics and in gravitation theory. In electrodynamics, generally speaking, the forces do depend on velocity (cf. eq. 37.1), and the following proof of the invariance of Newton's equations does not apply to these forces.

Consider a system of N interacting points, with coordinates in the moving system S′ given by x'_i, y'_i, z'_i $(i = 1, 2, \ldots, N)$. We shall write down Newton's equation of motion for one of these points, e.g., the first

$$m_1 \frac{d^2x'_1}{dt'^2} = F'_{1x} = -\frac{\partial V'_1}{\partial x'_1} \qquad m_1 \frac{d^2y'_1}{dt'^2} = F'_{1y} = -\frac{\partial V'_1}{\partial y'_1}$$

$$m_1 \frac{d^2z'_1}{dt'^2} = F'_{1z} = -\frac{\partial V'_1}{\partial z'_1} \tag{63.13}$$

where V'_1 is the potential energy of this point, dependent only on the distances of this point from the other points, i.e.

$$V'_1 = V'_1(r'_{12}, r'_{13}, \ldots, r'_{1N}) \tag{63.14}$$

where

$$r'_{1i} = \sqrt{(x'_1 - x'_i)^2 + (y'_1 - y'_i)^2 + (z'_1 - z'_i)^2}$$

is the distance between the first and the i^{th} point. Let us now transfer to a system S at rest. By (63.4), the distances r'_{1i} in (63.14) do not change. Hence, the function V'_1 does not change, i.e., we have

$$V'_1 = V'_1(r'_{12}, r'_{13}, \ldots, r'_{1N}) = V'_1(r_{12}, r_{13}, \ldots r_{1N}) = V_1 \tag{63.15}$$

where

$$r_{1i} = \sqrt{(x_1 - x_i)^2 + (y_1 - y_i)^2 + (z_1 - z_i)^2}$$

is the distance between the first and the i^{th} point in S. It follows also from the transformation formulas that

$$\frac{\partial V'_1}{\partial x'_1} = \frac{\partial V_1}{\partial x_1} \qquad \frac{\partial V'_1}{\partial y'_1} = \frac{\partial V_1}{\partial y_1} \qquad \frac{\partial V'_1}{\partial z'_1} = \frac{\partial V_1}{\partial z_1} \tag{63.16}$$

In classical mechanics, the mass of a particle is assumed to be constant, independent of the particle velocity. Taking (63.12) and (63.16) into account, we may rewrite (63.13) in the form

$$m_1 \frac{d^2x_1}{dt^2} = -\frac{\partial V_1}{\partial x_1} = F_{1x} \qquad m_1 \frac{d^2y_1}{dt^2} = -\frac{\partial V_1}{\partial y_1} = F_{1y}$$

$$m_1 \frac{d^2z_1}{dt^2} = -\frac{\partial V_1}{\partial z_1} = F_{1z} \tag{63.17}$$

These are Newton's equations in the coordinate system at rest, S. Comparison of (63.17) and (63.13) shows that Newton's equations are invariant under the Galilean transformation. We must stress once more that this proof holds only for forces which depend solely on the distances between the interacting bodies.

Principle of Relativity in Classical Mechanics. Newton's laws are formulated in the same way in all inertial systems. Newton's equations of motion have the same form in all inertial systems. Consequently, mechanical phenomena follow completely identical courses in all inertial systems, and all inertial systems are equivalent for the description of mechanical phenomena. Hence, uniform rectilinear motion of a coordinate system has no effect on the course of the mechanical processes taking place in that system. Therefore, no mechanical experiments carried out in a system can provide any evidence of the state of motion of that system with respect to other inertial systems. This is the *principle of relativity of classical mechanics.*

The mathematical formulation of this principle may be given as follows: the laws of mechanics are invariant under the Galilean transformation.

§64. The Principle of Relativity in the Special Theory of Relativity

The principle of relativity of classical mechanics states that it is impossible to detect uniform rectilinear motion by means of mechanical tests, but it makes no assertion that it is impossible to do so by means of other phenomena, e.g., electrodynamic phenomena. This possibility might present itself if the laws of some phenomenon were dependent on the velocity of an inertial coordinate system. In that case, once we had established the dependence, we could use this phenomenon to demonstrate the uniform rectilinear motion of a coordinate system.

Such attempts to demonstrate uniform rectilinear motion were first made using electrodynamic phenomena. In Chapter 12, we have described how attempts have been made to demonstrate a state of uniform motion by means of experiments using light. However, these attempts were unsuccessful. Many other experiments have been carried out that were intended to demonstrate the uniform rectilinear motion of a coordinate system, but they all gave negative results. Hence, the principle of relativity of classical mechanics had to be extended to all phenomena.

The laws of nature are, therefore, formulated in the same manner in all inertial systems. Consequently, all phenomena take place in a completely identical manner in all inertial coordinate systems, and all inertial coordinate systems are equivalent for the description of natural phenomena.

Hence, it is impossible to carry out experiments in one inertial system to obtain evidence of its state of motion with respect to other inertial systems. This is the *principle of relativity of the special theory of relativity.*

The principle of relativity is a postulate, i.e., an assertion adopted as a starting point without proof. The experimental facts which have been cited confirm the correctness of this principle. A direct proof of this principle is impossible.

The statement that the velocity of light is constant is incompatible with the Galilean transformation, yet all the experimental evidence forces us to accept it. Hence, we must abandon the Galilean transformation, and investigate other transformations, under which the velocity of light is invariant, i.e., the Lorentz transformation. The Galilean transformation is obtained from the Lorentz transformation in the special case of velocities much less than the velocity of light, which is always the case with the velocities encountered in everyday experience. In this case, there is, practically speaking, no difference between the Lorentz transformation and Galileo's, which is the reason why it took so long for the inaccuracies of the Galilean transformation to come to light.

CHAPTER 14

The Lorentz Transformation and Its Kinematic Corollaries

§65. Derivation of the Lorentz Transformation

In the preceding two chapters we have considered the experimental observations and concepts leading to the postulate that the velocity of light is constant, and to the principle of relativity. The entire special theory of relativity is based on these two propositions.

Synchronization of Clocks. Before proceeding to the derivation of the Lorentz transformation, we shall discuss the synchronization of clocks at different points in space. Since the velocity of light is constant, we can use light signals for synchronizing clocks. Completely identical clocks are placed at every point of a system under consideration. The identity of behavior of these clocks means that if light signals are radiated from a given point with a time interval Δt between them, the time interval between the instants when the signals are received at another point is, according to the clock at that point, also equal to Δt. This is the meaning of the assertion that the clocks are synchronized. Synchronization of the readings of the clocks is carried out as follows. A light signal is emitted from a point A when the clock at A shows time t_A. Let the point B be a distance r_{AB} from A. At the instant when the signal reaches the clock at B, this clock shows, by definition, the time

$$t_B = t_A + \frac{r_{AB}}{c} \tag{65.1}$$

This definition is based on the postulate of the constancy of the velocity of light. If a clock A is synchronized with a clock B, and the clock B is synchronized with a clock D, then the situation is equivalent to the direct synchronization of the clock A with the clock D. It must not be thought that the presence of clocks at every point of the coordinate system plays any significant part in this synchronization procedure. In principle, we could measure the time at any point of a coordinate system using a single clock at some point in the system, but then it would be necessary to apply the calculation (65.1) every time. This makes the discussion more complicated and less clear, but there is no difference in principle. Neither should it be thought that light signals play any significant role. Generally speaking, the relationship between time at different points in space could be established by means other than light signals. If we know that an effect travels from a point A to a point B with a velocity u, then clock A and clock B could be synchronized using this effect, employing equation (65.1), and substituting u for c. We would simply have to be sure that in a given coordinate system, the effect does actually travel from A to B with the constant velocity u. It is easy to see that using this process for the synchronization, we obtain exactly the same result as using light signals. However, light signals are used, because the laws of the propagation of light *in vacuo* have been studied in great detail, and formulated in the postulate c = const which lies at the root of the theory of relativity.

Mutual Orientation of Coordinate Systems. As in the case of the derivation of the Galilean transformation, we shall limit our discussion of the transformation of coordinates to the case shown in Fig. 75. A rotation of the axes and a displacement of the origin are, in the absence of motion, purely geometrical processes, effected by geometrical means.

Linear Nature of the Transformation. Space is homogeneous and the origin of a coordinate system is not distinguished by any feature. Moreover, the transformation must not distinguish one inertial system from another. Hence, it follows that the transformations must be linear.

Transformations for y and z. The linear transformations for these coordinates have the form

$$\left.\begin{aligned} y' &= a_1x + a_2y + a_3z + a_4t \\ z' &= b_1x + b_2y + b_3z + b_4t \end{aligned}\right\} \tag{65.2}$$

We assume that the y' axis is parallel to the y axis and the z' axis parallel to the z axis. Since the x' axis always coincides with the x axis, we see that $y = 0$ always implies $y' = 0$, and $z = 0$ always implies $z' = 0$. Hence, in the transformations (65.2), we must put

$$a_1 = a_3 = b_1 = b_2 = 0 \qquad a_4 = b_4 = 0 \tag{65.3}$$

giving

$$y' = ay \qquad z' = az \tag{65.4}$$

where we have assumed that since the axes are equivalent with respect to motion, the coefficients in the transformation must be identical: $a_2 = b_3 = a$.

The coefficient a in (65.4) indicates how many times greater is the scale in the S′ system of coordinates compared with S.

Let us rewrite equation (65.4) in the form

$$y = \frac{1}{a} y' \qquad z = \frac{1}{a} z' \tag{65.5}$$

The coefficient $1/a$ in equation (65.5) shows how many times smaller is the scale in S than in S′.

Since, according to the principle of relativity, both coordinate systems are equivalent, the change of scale from S to S′ must be the same as the change of scale from S′ to S. Hence, in (65.4) and (65.5) we must put $a = 1/a$, from which it follows that

$$a = 1$$

Hence, the transformations for y and z are

$$y' = y \qquad z' = z \tag{65.6}$$

Transformations for x and t. Since the variables y and z are transformed separately, the variables x and t can only be related, in a linear transformation, to each other. The origin of the moving system S′ is described in the coordinates of S by

$$x = vt \tag{65.7}$$

and in S′ by

$$x' = 0 \tag{65.8}$$

Since the transformation is linear, it must follow that

$$x' = \alpha(x - vt) \tag{65.9}$$

where α is a coefficient of proportionality, which is to be determined.

We shall now consider S′ to be at rest, and S to be moving. Then the origin of S is given in the S′ coordinates by

$$x' = -vt' \tag{65.10}$$

because in the S′ coordinate system, the S system is moving in the negative x direction. The origin of S is given in the S coordinates by

$$x = 0 \tag{65.11}$$

Hence, taking S′ as the system at rest, in place of (65.9) we obtain

$$x = \alpha'(x' + vt') \tag{65.12}$$

where α' is the coefficient of proportionality. We shall now show that by the principle of relativity

$$\alpha = \alpha' \tag{65.13}$$

Consider a rod at rest in S′ of length l in S′. This means that the ends of the rod in S′ have the coordinates

$$x'_1 = 0 \qquad x'_2 = l \tag{65.14}$$

In S, this rod moves with a velocity v. We take the length of a moving rod to be the distance between two points in S which coincide with the ends of the moving rod. We note the positions of the ends of the rod at a time t_0. Let these be x_1 and x_2. Applying equations (65.14) and (65.9), we obtain

$$x'_1 = \alpha(x_1 - vt_0) \qquad x'_2 = \alpha(x_2 - vt_0) \tag{65.15}$$

Hence, the length of the moving rod in S is equal to

$$x_2 - x_1 = \frac{x'_2 - x'_1}{\alpha} = \frac{l}{\alpha} \tag{65.16}$$

Now let the same rod be at rest in S. Then its length in S is l and, hence, the coordinates of its ends are

$$x_1 = 0 \qquad x_2 = l \tag{65.17}$$

Then, from the point of view of S′ (taken to be at rest), the rod is moving with velocity $-v$. To measure its length with respect to S′, we must note the positions of its ends at some instant t'_0 measured in S′. Using equations (65.17) and (65.12), we have

$$x_1 = \alpha'(x'_1 + vt'_0) \qquad x_2 = \alpha'(x'_2 + vt'_0) \tag{65.18}$$

Hence, the length of the moving rod in S′ (taken to be at rest) is equal to

$$x'_2 - x'_1 = \frac{x_2 - x_1}{\alpha'} = \frac{l}{\alpha'} \tag{65.19}$$

According to the principle of relativity, both systems are equivalent, and therefore, the length of the same rod moving with the same velocity will be the same in both cases. Hence, (65.13) follows from (65.16) and (65.19).

We shall now use the postulate that c = const. Let the origins of the two systems coincide at $t = t' = 0$. At that instant, let a light signal be emitted from the origin. The propagation of this signal in S and in S′ is described by

$$x' = ct' \qquad x = ct \tag{65.20}$$

which take into account the fact that the velocity of light is c in both systems.

Substituting (65.20) in (65.9) and (65.12), and using (65.13), we obtain

$$ct' = \alpha t(c - v) \qquad ct = \alpha t'(c + v) \tag{65.21}$$

thus, multiplying term by term, it follows that

$$\alpha = \frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{65.22}$$

From (65.12) we have

$$vt' = \frac{x}{\alpha} - x' = \frac{x}{\alpha} - \alpha(x - vt) = x\left(\frac{1}{\alpha} - \alpha\right) + \alpha vt$$

using equation (65.9). Consequently

$$t' = \alpha\left\{t + \frac{x}{v}\left(\frac{1}{\alpha^2} - 1\right)\right\} \tag{65.23}$$

where α is given by (65.22).

Form of the Lorentz Transformation. From the above, it can be seen that the Lorentz transformation takes the form

$$\left.\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} \\ y' &= y \\ z' &= z \\ t' &= \frac{t - \dfrac{v}{c^2}x}{\sqrt{1 - \dfrac{v^2}{c^2}}} \end{aligned}\right\} \tag{65.24}$$

By the principle of relativity, the inverse transformation must be of the same form, but with the sign of the velocity reversed. Hence, we may write

$$\left.\begin{aligned} x &= \frac{x' + vt'}{\sqrt{1 - \dfrac{v^2}{c^2}}} \\ y &= y' \\ z &= z' \\ t &= \frac{t' + \dfrac{v}{c^2}x'}{\sqrt{1 - \dfrac{v^2}{c^2}}} \end{aligned}\right\} \tag{65.25}$$

The transition from (65.24) to (65.25) may be effected without using the principle of relativity, since equations (65.25) may be obtained by solving equations (65.24) for the S coordinates.

Form of the Galilean Transformation as a Special Case of the Lorentz Transformation. In the limiting case of velocities much lower than the velocity of light, we may ignore terms of the order of

$$\frac{v}{c} \ll 1 \tag{65.26}$$

in the Lorentz transformation. Then (65.24) assumes the form of the Galilean transformation (63.1). Thus, at low velocities, the differences between the Lorentz transformation and the Galilean transformation are negligible, and that is why the inaccuracies of the Galilean transformation have remained undetected for so long.

Invariance of the Interval. The interval between two points x_1, y_1, z_1, t_1 and x_2, y_2, z_2, t_2 in the theory of relativity is a quantity, the square of which is given by

$$s^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 \tag{65.27}$$

This interval is invariant under the Lorentz transformation. To verify this, we must transform (65.27) into the S′ coordinates, using equations (65.25). We then have

$$x_2 - x_1 = \frac{x'_2 - x'_1 + v(t'_2 - t'_1)}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$y_2 - y_1 = y'_2 - y'_1$$

$$z_2 - z_1 = z'_2 - z'_1$$

$$t_2 - t_1 = \frac{t'_2 - t'_1 + \frac{v}{c^2}(x'_2 - x'_1)}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Substituting these equations in (65.27), we find that

$$\begin{aligned} s^2 &= (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 \\ &= (x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - c^2(t'_2 - t'_1)^2 = s'^2 \end{aligned} \tag{65.28}$$

This proves that the value of the square of an interval is the same in all coordinate systems, i.e., it is invariant.

If we consider two points an infinitesimal distance apart, then equation (65.28) proves the invariance of the square of the differential of the interval

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2\,dt^2 = \text{inv} \tag{65.29}$$

Invariance of the Wave Equation. The scalar and vector potentials which describe the propagation of waves in empty space obey a wave equation of the form

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2} = 0 \tag{65.30}$$

We shall show that this equation is invariant under the Lorentz transformation. To write it in the moving S′ system, we must replace the variables in (65.30), using the substitutions of (65.25) and expressing the derivatives with respect to the S variables in terms of the derivatives with respect to the S′ variables, using (65.24). We then have

$$\frac{\partial\Phi}{\partial x} = \frac{\partial\Phi}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial\Phi}{\partial y'}\frac{\partial y'}{\partial x} + \frac{\partial\Phi}{\partial z'}\frac{\partial z'}{\partial x} + \frac{\partial\Phi}{\partial t'}\frac{\partial t'}{\partial x} \tag{65.31}$$

From (65.24), it follows that

$$\frac{\partial x'}{\partial x} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \frac{\partial y'}{\partial x} = 0$$

$$\frac{\partial z'}{\partial x} = 0 \qquad \frac{\partial t'}{\partial x} = \frac{-\frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Hence, instead of (65.31), we have

$$\frac{\partial\Phi}{\partial x} = \frac{\partial\Phi}{\partial x'}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{\partial\Phi}{\partial t'}\frac{\frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Differentiating again, we obtain

$$\frac{\partial^2\Phi}{\partial x^2} = \frac{\partial^2\Phi}{\partial x'^2}\frac{1}{1 - \frac{v^2}{c^2}} - 2\frac{\partial^2\Phi}{\partial x'\,\partial t'}\frac{\frac{v}{c^2}}{1 - \frac{v^2}{c^2}} + \frac{\partial^2\Phi}{\partial t'^2}\frac{\frac{v^2}{c^4}}{1 - \frac{v^2}{c^2}}$$

The derivatives with respect to the other variables are found in exactly the same way

$$\frac{\partial^2\Phi}{\partial y^2} = \frac{\partial^2\Phi}{\partial y'^2} \qquad \frac{\partial^2\Phi}{\partial z^2} = \frac{\partial^2\Phi}{\partial z'^2}$$

$$\frac{\partial^2\Phi}{\partial t^2} = \frac{\partial^2\Phi}{\partial x'^2}\frac{v^2}{1 - \frac{v^2}{c^2}} - 2\frac{\partial^2\Phi}{\partial x'\,\partial t'}\frac{v}{1 - \frac{v^2}{c^2}} + \frac{\partial^2\Phi}{\partial t'^2}\frac{1}{1 - \frac{v^2}{c^2}}$$

Substituting these expressions in equation (65.30), we find

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t^2} = \frac{\partial^2\Phi}{\partial x'^2} + \frac{\partial^2\Phi}{\partial y'^2} + \frac{\partial^2\Phi}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2\Phi}{\partial t'^2} = 0$$

which proves that the wave equation is invariant under the Lorentz transformation.

Relativistic Invariance of the Laws of Nature. The Lorentz transformation is the mathematical expression of the postulate that the velocity of light is constant, and of the principle of relativity. It shows how the coordinates and time of an inertial reference system are interrelated. The principle of relativity may therefore be formulated as follows: *the laws of nature are invariant under the Lorentz transformation.*

If there were a law of nature which was not invariant under the Lorentz transformation, it would be formulated differently in different inertial coordinate systems. Then, from the way the law was formulated in a given system, it would be possible to make some deductions as to the state of motion of that system, or, in other words, inertial coordinate systems would not be equivalent. But this contradicts the principle of relativity; hence, the laws of nature must be invariant under the Lorentz transformation, or, as we say, they must be *relativistically invariant.*

Heuristic Importance of the Requirement of Relativistic Invariance. Relativistic invariance is of great heuristic importance. If a theory of some phenomenon is formulated, we must check first of all whether it is relativistically invariant. If it is found that the theory is not relativistically invariant, then it is basically wrong. At best it can only be an approximate theory which remains true only under certain conditions, and with a certain degree of precision. An investigation of the question of the relativistic invariance of a given approximation may help to establish the limits and conditions of its application.

Another aspect of the heuristic importance of the requirement of relativistic invariance is this: if it is necessary to establish the course of some process in a moving frame of reference, then it is sufficient to investigate its course in a frame of reference at rest. The process in the moving frame of reference may be found quite simply by applying the Lorentz transformation from the frame of reference at rest to the moving frame of reference.

§66. Length of a Moving Body

Determination of the Length of a Moving Body. The length of a moving body is the distance between the points, in a frame of reference at rest,

which coincide with the ends of the moving body at some instant of time reckoned by a clock in that frame of reference. It must be stressed that we are discussing the positions of the ends of the moving body at the same instant of time according to clocks in the frame of reference at rest. We must remember that the concept of simultaneity is relative: what is simultaneous in one frame of reference is not simultaneous in another. Therefore, whenever we use the concept of simultaneity, we must indicate the frame of reference to which it refers.

This concept of the length of a moving body has been also used in classical mechanics, where it has been assumed that different inertial coordinate systems are related by the Galilean transformation. Equation (63.4) shows that length is invariant under the Galilean transformation, i.e., the length of a body is constant in all frames of reference, irrespective of the state of motion of the body with respect to these frames of reference. Therefore, in classical mechanics, we may speak generally about the length of a body, without indicating the coordinate system we are considering.

In §65 it has been shown that inertial systems are related to one another, not by the Galilean transformation, but by the Lorentz transformation. Therefore, we must investigate the length of a body from the point of view of the Lorentz transformation.

Expression for the Contraction of the Length of a Moving Body. Let a rod of length l be at rest in a frame of reference S′, lying along the x' axis. The coordinates of the ends of the rod are x'_1 and x'_2. By definition, the length of the rod is

$$x'_2 - x'_1 = l \tag{66.1}$$

We note that the length of the rod l is written without a prime in this case, since it is the length of the rod in the frame of reference in which it is at rest, i.e., it is the length of the rod at rest, or its rest length.

We note the positions of the ends of the rod in a frame of reference S at a time t_0. The rod is moving with velocity v with respect to S. By the Lorentz transformation, we have

$$x'_1 = \frac{x_1 - vt_0}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad x'_2 = \frac{x_2 - vt_0}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{66.2}$$

Hence

$$l = x'_2 - x'_1 = \frac{x_2 - x_1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{l'}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{66.3}$$

where we are introducing the notation for the length of the rod in S

$$l' = x_2 - x_1 \tag{66.4}$$

Rewriting (66.3) in the form

$$l' = l\sqrt{1 - \frac{v^2}{c^2}} \tag{66.5}$$

we observe that the length of a moving rod, aligned along its direction of motion, is less than its rest length. Of course, if we repeat this analysis from the point of view of S′ (assumed to be at rest), then, by the principle of relativity, we must obtain the same expression. This concept has been used in the proof of (65.13).

If the rod length is perpendicular to the direction of motion, e.g., if it lies along the y' or z' axis, then we conclude from (65.24) that, in this case, there is no change in the rod length.

Change of Shape of Moving Bodies. The dimensions of bodies suffer contraction in the direction of motion, in accordance with (66.5), but remain unchanged in directions perpendicular to the motion. A body is, therefore, "flattened" in the direction of motion. If we record, at some instant of time in S, the positions of all points of a moving body, we obtain, as it were, a "cast" of the moving body. The shape of the cast is taken, by definition, to be the shape of the moving body. The shape of the cast is not the same as the shape of the body from which the cast is taken, when this body is at rest with respect to the cast; the cast is flattened compared with the body. This is the essence of the assertion that bodies become flattened in the direction of motion (the Fitzgerald contraction). This effect is a real effect in the sense discussed above.

Equation (66.5) shows that the length of a body is not invariant under the Lorentz transformation. Hence, it is impossible to say that a body has a given length without indicating the state of motion of the body. Usually, we take the dimensions of a body to mean the dimensions of that body at rest.

Estimate of the Fitzgerald Contraction. The velocities of material bodies are usually much less than the velocity of light, i.e., $(v/c) \ll 1$. Hence, we may put equation (66.5) in the form

$$l' = l\left(1 - \frac{1}{2}\frac{v^2}{c^2}\right) \tag{66.6}$$

with accuracy to the first order in v^2/c^2. The relative contraction in length then equals

$$\frac{\Delta l}{l} = \frac{l' - l}{l} = -\frac{1}{2}\frac{v^2}{c^2} \tag{66.7}$$

For velocities of the order of 10 km/sec, $v^2/c^2 \approx 10^{-8}$. Therefore, under ordinary conditions, the relative contraction is very small. For example, at the stated velocities, a length of 1 m suffers a contraction of the order of 10^{-6} cm.

§67. Rate of Moving Clocks. Proper Time

Slowing Down of Moving Clocks. Suppose that, at the same point in a moving (S′) system of coordinates, e.g., at the point x'_0 on the x' axis, two events occur at times t'_1 and t'_2. The time interval between them in the (S′) system is $\Delta t' = t'_1 - t'_2$, and in a system at rest (S) it is $\Delta t = t_2 - t_1$.

Applying the Lorentz transformation, we may write

$$t_1 = \frac{t'_1 + \dfrac{v}{c^2}x'_0}{\sqrt{1 - \dfrac{v^2}{c^2}}} \qquad t_2 = \frac{t'_2 + \dfrac{v}{c^2}x'_0}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{67.1}$$

Hence, it follows that

$$\Delta t = t_2 - t_1 = \frac{t'_2 - t'_1}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \frac{\Delta t'}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{67.2}$$

The time interval between the two events, measured by the moving clock

$$\Delta t' = \Delta t\sqrt{1 - \frac{v^2}{c^2}} \tag{67.3}$$

is less than the time interval Δt between the same events, measured by the clock at rest. This means that the rate of a moving clock is slower than that of a clock at rest.

At first glance, this assertion seems to contradict the principle of relativity, since a moving clock may be considered to be at rest. However, this contradiction is only apparent. The fact is that in equation (67.3) we compare time at the same point in S′ with time at different points in S. Hence, to apply the principle of relativity we would have to compare time at different points in S′ with time at the same point in S. Let us make this comparison. Let us suppose that at some point in S, e.g., x_0 on the x axis, two successive events occur at times t_1 and t_2. The time interval between these events is $\Delta t = t_2 - t_1$. In the S′ system (taken to be at rest), these events occur at different points at times t'_1 and t'_2. From equation (65.24), we have

$$t'_1 = \frac{t_1 - \frac{v}{c^2} x_0}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad t'_2 = \frac{t_2 - \frac{v}{c^2} x_0}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{67.4}$$

Hence, it follows that

$$\Delta t' = t'_2 - t'_1 = \frac{t_2 - t_1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{67.5}$$

but now $\Delta t'$ is the time interval between the events in a coordinate system at rest and Δt is in a moving coordinate system. Thus, the meaning of (67.5) is the same as the meaning of (67.2), and there is no contradiction of the principle of relativity.

Proper Time. The time measured by a clock fixed to a moving point is called the *proper time* of that point. Let us consider equation (67.3) for an infinitesimal time interval. We have

$$d\tau = dt \sqrt{1 - \frac{v^2}{c^2}} \tag{67.6}$$

where $d\tau$ is the differential of the proper time of the moving point, and dt is the differential of the time of an inertial coordinate system in which this point is moving, at a given instant, with velocity v.

Equation (65.28) shows that the interval is invariant. Remembering that

$$dx^2 + dy^2 + dz^2 = d\mathbf{r}^2 \tag{67.7}$$

is the three-dimensional distance in the system of coordinates under discussion, we may rewrite (65.29)

$$\frac{ds}{i} = c\, dt \sqrt{1 - \frac{1}{c^2}\left(\frac{d\mathbf{r}}{dt}\right)^2} = c\, dt \sqrt{1 - \frac{v^2}{c^2}} \tag{67.8}$$

where two successive positions of the moving point are taken as our two events. It is clear that

$$\left(\frac{d\mathbf{r}}{dt}\right)^2 = v^2 \tag{67.9}$$

is the square of the velocity of the moving point. Comparison of (67.6) and (67.8) shows that the proper time is proportional to the interval

$$d\tau = \frac{ds}{ic} \tag{67.10}$$

The velocity of light is constant, and ds is invariant (as was proved in equation (65.29)). Hence, the proper time is invariant under the Lorentz transformation.

Experimental Confirmation of the Time Dilatation. This important effect is directly confirmed by experiments on the decay of μ-mesons. The majority of known elementary particles exist only for times of 10^{-6} to 10^{-10} seconds, and then they are transformed into other particles. The length of time for which a particle exists is called its *lifetime*. From the macroscopic point of view a time interval of 10^{-6} to 10^{-10} seconds is small, but from the point of view of atomic processes, it is a long interval (for example, the period of revolution of an electron about a nucleus of an atom is of the order of 10^{-15} sec.)

Among the elementary particles there are π-*mesons* (pi-mesons). There are positive π^+-mesons, negative π^--mesons and neutral π^0-mesons. A positively charged π^+-meson decays into a μ^+-meson (mu-meson) and a neutrino ν according to the scheme

$$\pi^+ \longrightarrow \mu^+ + \nu \tag{67.11}$$

A *neutrino* is a neutral particle of rest mass equal to zero, and moving at the velocity of light. After a certain interval of time, a μ^+-meson decays in flight into a positron e^+ and a neutrino and anti-neutrino

$$\mu^+ \longrightarrow e^+ + \nu + \bar{\nu} \tag{67.12}$$

A *positron* is a particle of mass equal to the electron mass, but with a positive charge. The scheme of the formation and decay of a μ-meson is shown in Fig. 76.

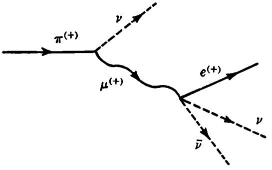


Fig. 76

There are various ways of recording charged particles. These enable us to measure the length of the path of a μ-meson between formation and decay, and to measure its velocity. Thus, we can determine its lifetime. If the time dilatation effect does occur, the lifetime of a μ-meson should increase with its velocity v

$$\tau_{\mu(+)} = \frac{\tau_{\mu(+)}{}^{(0)}}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{67.13}$$

where $\tau_{\mu(+)}{}^{(0)}$ is the lifetime measured by a clock fixed to the μ-meson, i.e., its proper lifetime; $\tau_{\mu(+)}$ is the lifetime according to a clock in the laboratory coordinate system. In the absence of time dilatation the path length l would be given, in terms of the velocity, by

$$l = \tau_{\mu(+)}{}^{(0)}v \tag{67.14}$$

i.e., it would be a linear function of the velocity, while in the case of time dilatation, l would be expressed by

$$l = \tau_{\mu(+)}{}^{(0)} \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{67.15}$$

Experiment confirms equation (67.15). The proper lifetime of a μ^+-meson equals $\tau_{\mu(+)}{}^{(0)} = 2 \times 10^{-6}$ sec. This proves the time dilatation effect for moving particles, and confirms the value given by equation (67.3).

§68. Simultaneity

Relative Nature of Simultaneity. Events are said to be simultaneous if they occur at the same time. Let us consider two events occurring at points x_1 and x_2 at a time t_0 in a coordinate system S at rest. In a moving system S′, these events occur at points x'_1 and x'_2 at instants t'_1 and t'_2 which may be found from equation (65.24)

$$t'_1 = \frac{t_0 - \frac{v}{c^2}x_1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad t'_2 = \frac{t_0 - \frac{v}{c^2}x_2}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{68.1}$$

Thus, in S′ these events are separated by a time interval

$$\Delta t' = t'_2 - t'_1 = \frac{\frac{v}{c^2}(x_1 - x_2)}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{68.2}$$

i.e., they are not simultaneous. Hence, if the statement that two events are simultaneous is to be meaningful, we must indicate to which coordinate system the statement refers. Events which are simultaneous in one coordinate system are not simultaneous in another.

The relative nature of simultaneity is illustrated in Fig. 77, which indicates the times (according to 65.24) shown by clocks at different points of a moving system at an instant $t = 0$ according to clocks of a system at rest.

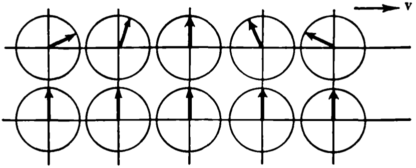


Fig. 77

Relative Nature of Simultaneity and Causality. It is clear from equation (68.2) that if $x_1 > x_2$, then if S′ is moving in the positive x direction, the relation $t'_2 > t'_1$ holds, but if S′ is moving in the opposite direction ($v < 0$), $t'_2 < t'_1$. Thus, the order in which the same two events occur is different in different coordinate systems. The question arises whether, if in some coordinate system the cause precedes the effect, in another coordinate system, the effect may precede the cause. To prevent such a state of affairs, we must assume that no effect on matter can be transmitted at a velocity greater than that of light. To prove this, we shall consider two events in a system at rest S. Let an event at point x_1 occurring at a time t_1 be the cause of an event at a point $x_2 > x_1$, which occurs at a time $t_2 > t_1$. The velocity of transmission of the "effect" from x_1 to x_2 is denoted by v_{infl}. Clearly, we have

$$\frac{x_2 - x_1}{v_{\text{infl}}} = t_2 - t_1 \tag{68.3}$$

In a moving system S′, these events occur at points x'_1 and x'_2 and times t'_1 and t'_2. Using (65.24), we may write

$$t'_2 - t'_1 = \frac{t_2 - t_1 - \frac{v}{c^2}(x_2 - x_1)}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{t_2 - t_1}{\sqrt{1 - \frac{v^2}{c^2}}}\left[1 - \frac{v}{c^2} v_{\text{infl}}\right] \tag{68.4}$$

eliminating $x_2 - x_1$ from the last equation by means of (68.3). Equation (68.4) shows that if

$$1 - \frac{v}{c^2} v_{\text{infl}} < 0 \tag{68.5}$$

then in S′ the effect occurs before the cause. But this is impossible. Hence, we must always have

$$1 - \frac{v}{c^2} v_{\text{infl}} \geqslant 0$$

i.e.

$$v_{\text{infl}} \leqslant \frac{c}{v} c \tag{68.6}$$

Since the Lorentz transformation permits values for v that are as close as we like to the speed of light, but not exceeding it (the transformation then ceases to be real), the requirement (68.6) must be written in the form

$$v_{\text{infl}} \leqslant c \tag{68.7}$$

Thus, the transmission of the effect from one point to another cannot occur with a velocity greater than the velocity of light. Under these conditions, the causality connection between events is absolute: there exists no coordinate system in which cause and effect can change places.

Space-Like and Time-Like Intervals. If a time interval between two events is denoted by t, and a space interval by l, then it follows from (65.28) that the square of

$$s^2 = l^2 - c^2t^2 \tag{68.8}$$

is invariant.

Let us suppose that in some coordinate system these events are not connected by any causal connection. Then, for these events

$$l > ct \tag{68.9}$$

or

$$s^2 > 0 \tag{68.10}$$

Since the interval is invariant, it follows that there will be no causal connection between these events in any other coordinate system. The converse result is also true, of course. If, in some coordinate system, events may, in principle, be connected by a causal connection, then, in principle, they may be connected by a causal connection in all other coordinate systems.

An interval for which

$$s^2 > 0 \tag{68.11}$$

is said to be *space-like*, and one for which

$$s^2 < 0 \tag{68.12}$$

is said to be *time-like*.

If an interval is space-like, then we can choose a coordinate system such that two events occur simultaneously at different points of space, and there is no coordinate system in which these two events will take place at the same point.

If an interval is time-like, then we can choose a coordinate system in which two events occur at the same point of space but at different times,

and there is no coordinate system in which these events will occur simultaneously.

A space-like or a time-like interval between two events is independent of the coordinate system. It is an invariant property of the two events.

In short, we may say that a causal connection can exist only between events separated by a time-like interval. There can be no causal connection between events separated by a space-like interval. This is due to the fact that the velocity of transmission of an effect cannot exceed the velocity of light.

§69. Addition of Velocities

Formula for the Addition of Velocities. Let the motion of a material point in a moving system of coordinates S′ be given by

$$x' = x'(t') \qquad y' = y'(t') \qquad z' = z'(t') \tag{69.1}$$

In a coordinate system at rest S, its motion will be given by

$$x = x(t) \qquad y = y(t) \qquad z = z(t) \tag{69.2}$$

which is obtained from (69.1) by the transformations (65.24). It is necessary to find the relationship between the components of the velocity of the point in S′

$$u'_x = \frac{dx'}{dt'} \qquad u'_y = \frac{dy'}{dt'} \qquad u_z = \frac{dz'}{dt'} \tag{69.3}$$

and in S

$$u_x = \frac{dx}{dt} \qquad u_y = \frac{dy}{dt} \qquad u_z = \frac{dz}{dt} \tag{69.4}$$

From (65.25), we have

$$dx = \frac{dx' + v\,dt'}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad dy = dy' \qquad dz = dz'$$

$$dt = \frac{dt' + \frac{v}{c^2}\,dx'}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{69.5}$$

Substituting the values of the differentials from (69.5) in (69.4), and using (69.3), we obtain

$$u_x = \frac{u'_x + v}{1 + \frac{vu'_x}{c^2}} \qquad u_y = \frac{\sqrt{1 - \frac{v^2}{c^2}}\,u'_y}{1 + \frac{vu'_x}{c^2}} \qquad u_z = \frac{\sqrt{1 - \frac{v^2}{c^2}}\,u'_z}{1 + \frac{vu'_x}{c^2}} \tag{69.6}$$

The expressions for the inverse transformation are obtained, as usual, from the principle of relativity, writing $-v$ for v.

From (69.6) it follows that the sum of two velocities cannot exceed the velocity of light. Let us assume that $u'_y = u'_z = 0$, $u'_x = c$. Then, from (69.6), we find

$$u_x = \frac{c + c}{1 + \frac{c\,c}{c^2}} = c \qquad u_y = 0 \qquad u_z = 0 \tag{69.7}$$

This result is completely natural, since the equations were obtained from the requirement that the velocity of light is constant.

Aberration. Consider a ray of light propagated along the y' axis in S′, i.e.

$$u'_x = 0 \qquad u'_y = c \qquad u'_z = 0 \tag{69.8}$$

In S, we obtain

$$u_x = v \qquad u_y = \sqrt{1 - \frac{v^2}{c^2}}\,c \qquad u_z = 0 \tag{69.9}$$

Hence, in S the ray makes an angle α with the y axis, given by

$$\tan\alpha = \frac{u_x}{u_y} = \frac{v}{c}\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{69.10}$$

Since $(v/c) \ll 1$

$$\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \approx 1 + \frac{1}{2}\frac{v^2}{c^2}$$

equation (69.10) may be written in the form

$$\tan\alpha \approx \frac{v}{c} \tag{69.11}$$

with precision to the third order in v/c. In this form, it is identical with equation (57.5) of the classical theory, but its meaning is different. In the classical theory, we distinguish between the cases of a moving source with an observer at rest and a moving observer with the source at rest. In the theory of relativity, there is only one case of relative motion of source and observer.

Interpretation of Fizeau's Experiment. The result of Fizeau's experiment follows naturally from the formula for the addition of velocities of the theory of relativity.

The velocity of light with respect to a medium at rest is c/n, where n is the refractive index of the medium. Taking the x axis in the direction of

motion of the medium, in a moving coordinate system S′, we have the following expressions for the velocity of light

$$u'_x = \frac{c}{n} \qquad u'_y = 0 \qquad u'_z = 0 \tag{69.12}$$

From equation (69.6), we can now find the components of the velocity of light in the coordinate system S with respect to which the medium is in motion

$$u_x = \frac{\dfrac{c}{n} \pm v}{1 \pm \dfrac{v}{cn}} \qquad u_y = 0 \qquad u_z = 0 \tag{69.13}$$

where the plus sign refers to the case when the directions of motion of light and medium are the same, and the minus sign to the case when they are opposite.

Since $(v/c) \ll 1$, we can rewrite expression (69.13)

$$u_x \approx \left(\frac{c}{n} \pm v\right)\left(1 \mp \frac{v}{cn}\right) \approx \frac{c}{n} \mp \frac{v}{n^2} \pm v = \frac{c}{n} \pm \left(1 - \frac{1}{n^2}\right) v \tag{69.14}$$

where terms of the order of v/c and higher have been dropped; taking into account (60.4) ($u' = c/n$), this expression agrees with equation (60.1). Thus, Fizeau's experiment confirms the formula for the addition of velocities.

PROBLEMS

1 What is the contraction of the diameter of the earth in the direction of its motion about the sun from the point of view of an observer at rest relative to the sun? The radius of the earth $r = 6.4 \times 10^3$ km, the velocity of the earth about the sun $v = 30$ km/sec.
Solution:

$$\Delta d = 2r\left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right) \approx r\frac{v^2}{c^2} = 6.4 \times 10^{-2}\text{ m} = 6.4\text{ cm}$$

2 What should the velocity of a body be for its dimensions in the direction of its motion to be contracted by a factor of 2?
Solution:

$$\frac{l}{2} = l\sqrt{1 - \frac{v^2}{c^2}} \qquad v = \frac{\sqrt{3}}{2}c$$

3 In a coordinate system in which a μ^+-meson is at rest, its lifetime is equal to $\tau_{\mu(+)}{}^{(0)} = 2 \times 10^{-6}$ sec. Calculate the lifetime of a μ^+-meson and the distance which it will travel, if its velocity is $v = 0.99$ c.

$$Answer:\quad \tau_{\mu}{}^{+} = \frac{\tau_{\mu(+)}{}^{(0)}}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{\tau_{\mu(+)}{}^{(0)}}{0,1} = 2 \times 10^{-5}\,\text{sec}$$

$$l = v\tau \approx 6 \times 10^3\,\text{m} = 6\,\text{km}$$

4 A train is moving at 100 km/hour. A man walks up the train in the direction of its motion at a velocity of 5 km/hour. Calculate the difference between the velocities of the man relative to the track calculated by the formulas for the addition of velocities of classical physics and the theory of relativity.
Solution:

$$u_{\text{rel}} = \frac{u' + v}{1 + \frac{vu'}{c^2}} \approx u' + v' - (u' + v)\frac{vu'}{c^2}$$

$$\Delta u = u_{\text{cl}} - u_{\text{rel}} = (u' + v)\frac{vu'}{c^2} \approx 4.85 \times 10^{-14}\,\text{Km/hr}$$

5 In a coordinate system moving with a velocity v, which is close to the velocity of light, two rays of light are emitted in the positive and negative y' directions (the y' axis is perpendicular to the direction of motion). In the moving system of coordinates, these rays are inclined to each other at an angle of 180°.

Determine the angle between the rays in a coordinate system at rest.
Solution: Using equations (69.6)

$$u'_{x1} = u'_{x2} = 0 \qquad u'_{y1} = c \qquad u'_{y2} = -c \qquad u'_{z1} = u'_{z2} = 0$$

Consequently

$$u_{x1} = v \qquad u_{x2} = v \qquad u_{y1} = c\sqrt{1 - \frac{v^2}{c^2}}$$

$$u_{y2} = -c\sqrt{1 - \frac{v^2}{c^2}} \qquad u_{z1} = u_{z2} = 0$$

If the angle between the rays is α, then

$$\tan\frac{\alpha}{2} = \frac{u_{y1}}{u_{x1}} = \frac{-u_{y2}}{u_{x2}} = \frac{c}{v}\sqrt{1 - \frac{v^2}{c^2}}$$

When $v \approx c$, this angle is very small, and the rays are almost parallel in the direction of the moving system of coordinates.

CHAPTER 15

Mathematical Apparatus of the Theory of Relativity

§70. Four-Dimensional Space

Three-Dimensional Space. It is well known that we live in three-dimensional space. This means that to describe the position of any point in space we must specify three numbers—the coordinates of that point in a chosen coordinate system. For example, in a rectangular Cartesian system these numbers are the x, y, z coordinates of the point. It is often convenient to denote these coordinates by the same letter but with different indices. Therefore, we shall use the notation

$$x_1 = x \qquad x_2 = y \qquad x_3 = z \tag{70.1}$$

Here, x_1, x_2, x_3 are the components of the radius vector $\mathbf{R}$, from the origin to a given point. We denote the unit vectors along the coordinate axes x_1, x_2, x_3 by $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, respectively. Then

$$\mathbf{R} = \mathbf{r}_1x_1 + \mathbf{r}_2x_2 + \mathbf{r}_3x_3 = \sum_{\alpha=1}^{3} \mathbf{r}_\alpha x_\alpha \tag{70.2}$$

Let us now consider a second coordinate system, the origin of which coincides with the first system but the axes do not coincide. We shall denote the unit vectors of this coordinate system by $\mathbf{r}'_1$, $\mathbf{r}'_2$, $\mathbf{r}'_3$, and the components of $\mathbf{R}$ in this system by x'_1, x'_2, x'_3. We then have

$$\mathbf{R} = \mathbf{r}'_1x'_1 + \mathbf{r}'_2x'_2 + \mathbf{r}'_3x'_3 = \sum_{\alpha=1}^{3} \mathbf{r}'_\alpha x'_\alpha \tag{70.3}$$

Since the coordinate system is orthogonal

$$\mathbf{r}_\alpha \cdot \mathbf{r}_\beta = \begin{cases} 0 & \text{for } \alpha \neq \beta \\ 1 & \text{for } \alpha = \beta \end{cases} \tag{70.4}$$

and similarly for the r′ system of coordinates. Multiplying scalarly both sides of (70.2) by $\mathbf{r}_\beta$ and using (70.4), we obtain the following expressions for the components of $\mathbf{R}$

$$x_\beta = \mathbf{R} \cdot \mathbf{r}_\beta \tag{70.5}$$

From (70.2) and (70.3), we have

$$\mathbf{R} = \sum_{\alpha=1}^{3} \mathbf{r}'_\alpha x'_\alpha = \sum_{\alpha=1}^{3} \mathbf{r}_\alpha x_\alpha \tag{70.6}$$

Multiplying both sides of this equation by $\mathbf{r}'_\beta$ and using (70.4), we find

$$x'_\beta = \sum_{\alpha=1}^{3} \mathbf{r}'_\beta \cdot \mathbf{r}_\alpha x_\alpha \tag{70.7}$$

We shall use the following notation for the scalar products of the unit vectors in different systems of coordinates

$$\mathbf{r}'_\beta \cdot \mathbf{r}_\alpha = a_{\beta\alpha} \tag{70.8}$$

Using this notation, the transformation (70.7) is now written as follows

$$x'_\beta = \sum_{\alpha=1}^{3} a_{\beta\alpha} x_\alpha \tag{70.9}$$

The inverse transformation is obtained from (70.6) by multiplying the two sides of that equation scalarly by $\mathbf{r}_\beta$. Thus, we obtain

$$x_\beta = \sum_{\alpha=1}^{3} \mathbf{r}_\beta \cdot \mathbf{r}'_\alpha x'_\alpha \tag{70.7a}$$

Using the notation of (70.8), we find

$$x_\beta = \sum_{\alpha=1}^{3} a_{\alpha\beta} x'_\alpha \tag{70.9a}$$

Thus, the components of the radius vector are transformed from one system of coordinates to another by linear transformations of the (70.9) type.

If we have any three quantities A_1, A_2, A_3, which are transformed from one system of coordinates to another in accordance with the transformation formula (70.9), i.e.

$$A'_\beta = \sum_{\alpha=1}^{3} a_{\beta\alpha} A_\alpha \qquad A_\beta = \sum_{\alpha=1}^{3} a_{\alpha\beta} A'_\alpha \tag{70.10}$$

the set of these three quantities A_1, A_2, A_3 is called a *three-dimensional*

vector, and the quantities themselves are called the components of the vector along the coordinate axes x_1, x_2, x_3.

It follows from (70.2), (70.3), and (70.4) that

$$R^2 = \sum_{\alpha=1}^{3} x_\alpha{}^2 = \sum_{\beta=1}^{3} x'_\beta{}^2 \tag{70.11}$$

i.e., the magnitude of the vector **R** is invariant under these transformations. Substituting (70.9) into (70.11), we have

$$\sum_{\alpha=1}^{3} x_\alpha{}^2 = \sum_{\alpha=1}^{3} \sum_{\beta=1}^{3} \sum_{\gamma=1}^{3} a_{\alpha\beta} x_\beta \, a_{\alpha\gamma} x_\gamma = \sum_{\beta,\gamma} x_\beta x_\gamma \sum_{\alpha=1}^{3} a_{\alpha\beta} a_{\alpha\gamma} \tag{70.12}$$

The left-hand and right-hand sides of (70.12) should be identical. Consequently, the terms $x_\beta x_\gamma$ should be absent on the right-hand side when $\beta \neq \gamma$, and the coefficients of the terms $x_\beta x_\gamma$ should be equal to unity when $\beta = \gamma$. Hence, we obtain the following properties of the coefficients $a_{\alpha\beta}$

$$\sum_{\alpha=1}^{3} a_{\alpha\beta} a_{\alpha\gamma} = \begin{cases} 0 & \text{for} \quad \beta \neq \gamma \\ 1 & \text{for} \quad \beta = \gamma \end{cases} \tag{70.13}$$

If we write down the equation analogous to (70.12) for the primed coordinates $x'_\alpha{}^2$, we obtain, in place of (70.13)

$$\sum_{\alpha=1}^{3} a_{\beta\alpha} a_{\gamma\alpha} = \begin{cases} 0 & \text{for} \quad \beta \neq \gamma \\ 1 & \text{for} \quad \beta = \gamma \end{cases} \tag{70.13a}$$

The coefficients $a_{\alpha\beta}$ may be written in the form of a matrix

$$a_{\alpha\beta} = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

The conditions (70.13) and (70.13a) mean that the product of different columns and rows in the matrix is equal to zero; a column or a row multiplied by itself is equal to unity. The concept of the product of columns and rows is defined by (70.13) and (70.13a). Equations (70.13) and (70.13a) are the conditions of the orthogonality of the transformation. They ensure that the absolute magnitude of a vector remains invariant under the transformation. For example, from (70.10), we have

$$\sum_{\beta=1}^{3} A'_\beta{}^2 = \sum_{\alpha,\beta,\gamma} a_{\beta\alpha} A_\alpha a_{\beta\gamma} A_\gamma = \sum_{\alpha,\gamma} A_\alpha A_\gamma \sum_{\beta} a_{\beta\alpha} a_{\beta\gamma} = \sum_{\beta=1}^{3} A_\beta{}^2 \tag{70.14}$$

i.e., the absolute magnitude of any vector is invariant under these transformations.

Hence, we may say that the absolute magnitude of a vector is an invariant under an orthogonal transformation of coordinates.

Space-Time. To describe a physical event completely, it is not sufficient to specify the space coordinates of the event; we must also specify time. Hence, a physical event in three-dimensional space is described by four numbers: three space coordinates x, y, z, and time t. A point in three-dimensional space is described by the set of three numbers x, y, z. By analogy, we may say that the four numbers x, y, z, t, describe a point in four-dimensional space, or *space-time.*

The transformation from one coordinate system to another in four-dimensional space is effected by the Lorentz transformation. This transformation is linear. To make it completely analogous to the transformation (70.9), we introduce a fourth variable describing time, using an imaginary coordinate, and introducing the following definitions

$$x_1 = x \qquad x_2 = y \qquad x_3 = z \qquad x_4 = ict \tag{70.15}$$

where $i = \sqrt{-1}$.

The Lorentz transformation is then written

$$\left.\begin{aligned}
x'_1 &= \frac{1}{\sqrt{1-\beta^2}}x_1 + 0\,x_2 + 0\,x_3 + \frac{i\beta}{\sqrt{1-\beta^2}}x_4 \\
x'_2 &= 0\,x_1 + 1\,x_2 + 0\,x_3 + 0\,x_4 \\
x'_3 &= 0\,x_1 + 0\,x_2 + 1\,x_3 + 0\,x_4 \\
x'_4 &= \frac{-i\beta}{\sqrt{1-\beta^2}}x_1 + 0\,x_2 + 0\,x_3 + \frac{1}{\sqrt{1-\beta^2}}x_4
\end{aligned}\right\} \tag{70.16}$$

where $\beta = v/c$.

The above transformation has the form

$$x'_\alpha = \sum_{\gamma=1}^{4} a_{\alpha\gamma}x_\gamma \tag{70.17}$$

where the coefficients $a_{\alpha\gamma}$ are defined by the matrix

$$a_{\alpha\gamma} = \left\| \begin{matrix} \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{i\beta}{\sqrt{1-\beta^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-i\beta}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{1}{\sqrt{1-\beta^2}} \end{matrix} \right\| \tag{70.18}$$

The inverse transformation, by analogy to (70.9a), is given by

$$x_\alpha = \sum_{\gamma=1}^{4} a_{\gamma\alpha}x'_\gamma \tag{70.17a}$$

It is easy to verify that the coefficients $a_{\alpha\gamma}$, defined by (70.18), satisfy the condition of orthogonality (70.13), which, in the case of four dimensions, takes the form

$$\sum_{\alpha=1}^{4} a_{\alpha\gamma}a_{\alpha\lambda} = \begin{cases} 0 & \text{for} \quad \gamma \neq \lambda \\ 1 & \text{for} \quad \gamma = \lambda \end{cases} \tag{70.19}$$

These coefficients also satisfy (70.13a).

Thus, in four-dimensional space, the transition from the coordinates of a point in one frame of reference to the coordinates in another frame of reference is effected by linear transformations of the form (70.17) and (70.17a), the coefficients of which are defined by the matrix (70.18).

§71. Four-Dimensional Vectors

Definition. A *four-dimensional vector* (or a four-vector) is an array of four quantities A_1, A_2, A_3, A_4, which are transformed from one coordinate system to another by formulas of the (70.17) and (70.17a) type, with the same coefficients $a_{\alpha\gamma}$, i.e.

$$A'_\alpha = \sum_{\gamma=1}^{4} a_{\alpha\gamma}A_\gamma \qquad A_\alpha = \sum_{\gamma=1}^{4} a_{\gamma\alpha}A'_\gamma \tag{71.1}$$

where $a_{\alpha\gamma}$ are defined by the matrix (70.18). Some of the values A_α may be complex.

Squaring both sides of (71.1), and summing over α, we find

$$\sum_{\alpha=1}^{4} A'^2_\alpha = \sum_{\alpha,\gamma,\mu} a_{\alpha\gamma}A_\gamma a_{\alpha\mu}A_\mu = \sum_{\gamma,\mu} A_\gamma A_\mu \sum_\alpha a_{\alpha\gamma}a_{\alpha\mu} = \sum_\alpha A_\alpha^2 \tag{71.2}$$

where the orthogonality condition (70.19) is used. The invariant

$$\sum_{\alpha=1}^{4} A_\alpha^2 = \text{inv} \tag{71.3}$$

is called the *square* of the four-dimensional vector.

The concept of four-dimensional vectors makes many calculations considerably easier, since, if an array of four values forms a vector, the law of transformation of these four values from one coordinate system to another is known.

Four-Dimensional Vector of a World Point. The set of coordinates of a point in four-dimensional space or *world point* is an example of a four-dimensional vector. The square of this vector equals

$$\sum_{\alpha=1}^{4} x_\alpha^2 = x_1^2 + x_2^2 + x_3^2 - c^2t^2 = s^2 \tag{71.4}$$

i.e., it is the square of the interval (65.27).

As was shown in (67.10), the differential of proper time $d\tau$ of a material point is invariant. Therefore, the set of derivatives with respect to $d\tau$ of the components of some four-dimensional vector also constitutes a four-dimensional vector, since differentiation with respect to an invariant does not change the properties of the transformation.

Four-Dimensional Velocity. Let us denote the derivatives of the coordinates of a world point with respect to proper time by

$$u_\alpha = \frac{dx_\alpha}{d\tau} \tag{71.5}$$

Then, the four-dimensional vector

$$u_1,\ u_2,\ u_3,\ u_4 \tag{71.6}$$

is called the *four-dimensional velocity*. Since

$$\begin{gathered} x_1 = x \qquad x_2 = y \qquad x_3 = z \qquad x_4 = ict \\ d\tau = dt\sqrt{1 - \frac{u^2}{c^2}} \end{gathered} \tag{71.7}$$

we may write the components of (71.6) in the form

$$\begin{aligned} u_1 &= \frac{u_x}{\sqrt{1 - \frac{u^2}{c^2}}} \qquad & u_2 &= \frac{u_y}{\sqrt{1 - \frac{u^2}{c^2}}} \\ u_3 &= \frac{u_z}{\sqrt{1 - \frac{u^2}{c^2}}} \qquad & u_4 &= \frac{ic}{\sqrt{1 - \frac{u^2}{c^2}}} \end{aligned} \tag{71.8}$$

where u_x, u_y, u_z are the components of the three-dimensional velocity; $u^2 = u_x^2 + u_y^2 + u_z^2$.

The square of the four-dimensional velocity equals

$$\sum_{\alpha=1}^{4} u_\alpha^2 = -c^2 \tag{71.9}$$

Transformation of the components of the four-dimensional velocity from one coordinate system to another is effected by equations (71.1)

$$u'_\alpha = \sum_\gamma a_{\alpha\gamma} u_\gamma \tag{71.10}$$

Using the values of $a_{\alpha\gamma}$ in (70.18), we obtain

$$\begin{gathered} u'_1 = \frac{u_1 + i\beta u_4}{\sqrt{1 - \beta^2}} \\ u'_2 = u_2 \qquad u'_3 = u_3 \\ u_4 = \frac{-i\beta u_1 + u_4}{\sqrt{1 - \beta^2}} \end{gathered} \tag{71.10a}$$

To go over to three-dimensional symbols, we substitute the expressions of (71.8) in (71.10a). This gives

$$\frac{u'_x}{\sqrt{1-\frac{u'^2}{c^2}}} = \frac{1}{\sqrt{1-\beta^2}} \frac{u_x - v}{\sqrt{1-\frac{u^2}{c^2}}} \tag{71.11a}$$

$$\frac{u'_y}{\sqrt{1-\frac{u'^2}{c^2}}} = \frac{u_y}{\sqrt{1-\frac{u^2}{c^2}}} \qquad \frac{u'_z}{\sqrt{1-\frac{u'^2}{c^2}}} = \frac{u_z}{\sqrt{1-\frac{u^2}{c^2}}} \tag{71.11b}$$

$$\frac{1}{\sqrt{1-\frac{u'^2}{c^2}}} = \frac{1}{\sqrt{1-\beta^2}} \frac{1-\frac{u_x v}{c^2}}{\sqrt{1-\frac{u^2}{c^2}}} \tag{71.11c}$$

where $\beta = v/c$; v is the relative velocity of the coordinate systems.

We may rewrite (71.11a) and (71.11b), using (71.11c), in the form

$$u'_x = \frac{u_x - v}{1-\frac{vu_x}{c^2}} \qquad u'_y = \frac{\sqrt{1-\beta^2}u_y}{1-\frac{vu_x}{c^2}} \qquad u'_z = \frac{\sqrt{1-\beta^2}u_z}{1-\frac{vu_x}{c^2}} \tag{71.12}$$

i.e., we obtain equation (69.6) for the addition of velocities. Thus, once we have established that (71.6) is a four-dimensional vector, we immediately obtain the correct formulas for transformation from one coordinate system to another. This demonstrates the convenience of introducing the concept of four-dimensional vectors.

Four-Dimensional Acceleration. When we differentiate the components of a four-dimensional vector (71.6) with respect to invariant proper time $d\tau$, we must once again obtain a four-dimensional vector with the components

$$b_\alpha = \frac{du_\alpha}{d\tau} \tag{71.13}$$

The four-dimensional vector

$$b_1, b_2, b_3, b_4 \tag{71.14}$$

is called the *four-dimensional acceleration*. The components of the four-dimensional acceleration can be expressed in three-dimensional terms

$$b_1 = \frac{du_1}{d\tau} = \frac{d}{dt}\left(\frac{u_x}{\sqrt{1-\frac{u^2}{c^2}}}\right)\frac{dt}{d\tau} = \frac{\dot{u}_x}{1-\frac{u^2}{c^2}} + \frac{u_x(\mathbf{u}\cdot\dot{\mathbf{u}})}{c^2\left(1-\frac{u^2}{c^2}\right)^2} \tag{71.15}$$

the expressions for b_2 and b_3 are analogous

$$b_4 = \frac{du_4}{d\tau} = \frac{d}{dt}\left(\frac{ic}{\sqrt{1-\frac{u^2}{c^2}}}\right)\frac{dt}{d\tau} = \frac{i}{c}\frac{\mathbf{u}\cdot\dot{\mathbf{u}}}{\left(1-\frac{u^2}{c^2}\right)^2} \tag{71.16}$$

The vector $\mathbf{u}$ in these expressions is the three-dimensional velocity vector; differentiation with respect to time t is denoted by a dot.

From (71.15) and (71.16) we may find immediately the invariant value of the square of the four-dimensional acceleration

$$\sum_{\alpha=1}^{4} b_\alpha^2 = \frac{\dot{\mathbf{u}}^2 - \left(\frac{\mathbf{u}}{c} \times \dot{\mathbf{u}}\right)^2}{\left(1 - \frac{u^2}{c^2}\right)^3} \tag{71.17}$$

Scalar Product of Four-Dimensional Vectors. By analogy to the three-dimensional case, the *scalar product* of two four-dimensional vectors A_1, A_2, A_3, A_4 and B_1, B_2, B_3, B_4 is defined by

$$(A, B) = \sum_{\alpha=1}^{4} A_\alpha B_\alpha$$

If the scalar product of two four-dimensional vectors equals zero, the vectors are said to be *orthogonal.* It is easy to see that the magnitude of the scalar product is invariant under orthogonal transformations of the coordinates.

Differentiating (71.9) with respect to $d\tau$, we obtain

$$\sum_{\alpha=1}^{4} u_\alpha b_\alpha = 0 \tag{71.18}$$

i.e., the four-dimensional acceleration is orthogonal to the four-dimensional velocity.

§72. Four-Dimensional Tensors

Three-Dimensional Tensors. Consider two vectors $\mathbf{A}$ and $\mathbf{B}$ with the components A_1, A_2, A_3 and B_1, B_2, B_3. Consider, further, the nine quantities $T_{\alpha\gamma}$ defined by the multiplication of the components A_α and B_γ according to the rule

$$T_{\alpha\gamma} = A_\alpha B_\gamma \tag{72.1}$$

The components A_α and B_γ are transformed according to (70.10). Hence, the transformation formulas for $T_{\alpha\gamma}$ are

$$T'_{\alpha\gamma} = A'_\alpha B'_\gamma = \sum_\mu a_{\alpha\mu} A_\mu \sum_\lambda a_{\gamma\lambda} B_\lambda = \sum_{\mu,\lambda} a_{\alpha\mu} a_{\gamma\lambda} A_\mu B_\lambda = \sum_{\mu,\lambda} a_{\alpha\mu} a_{\gamma\lambda} T_{\mu\lambda} \tag{72.2}$$

Consequently, $T_{\alpha\gamma}$ are transformed from one coordinate system to another according to the rule

$$T'_{\alpha\gamma} = \sum_{\mu,\lambda} a_{\alpha\mu} a_{\gamma\lambda} T_{\mu\lambda} \tag{72.3}$$

where the coefficients $a_{\alpha\gamma}$ have the same values as in the formulas for vector transformations. Thus, for each of their indices, $T_{\alpha\gamma}$ are transformed like vectors.

The array of nine values $T_{\alpha\gamma}$, which are transformed from one coordinate system to another by (72.3), are called a *three-dimensional tensor of the second rank.*

Four-Dimensional Tensors. A *four-dimensional tensor of the second rank* is defined as the array of 16 values ($4 \times 4 = 16$) which are transformed from one coordinate system to another according to the rule

$$T'_{\alpha\gamma} = \sum_{\mu,\lambda=1}^{4} a_{\alpha\mu}a_{\gamma\lambda}T_{\mu\lambda} \tag{72.4}$$

where the coefficients of the transformation are given by the matrix (70.18). In general, the array of values $T_{\alpha_1,\alpha_2,\ldots,\alpha_n}$, dependent on n indices, and transformed (from one coordinate system to another) with respect to every index like a vector, is called a tensor of the n^{th} rank. Hence, a vector is a tensor of the first rank, and a scalar may be considered to be a tensor of zero$^{\text{th}}$ rank.

Symmetric and Antisymmetric Tensors. A *symmetric tensor* is one in which

$$T_{\alpha\gamma} = T_{\gamma\alpha} \tag{72.5}$$

An *antisymmetric tensor* is defined by the condition

$$T_{\alpha\gamma} = -T_{\gamma\alpha} \tag{72.6}$$

Any tensor $T_{\alpha\gamma}$ may be written as the sum of a symmetric $T_{\alpha\gamma}{}^{s}$, and an antisymmetric tensor $T_{\alpha\gamma}{}^{a}$

$$T_{\alpha\gamma} = T_{\alpha\gamma}{}^{s} + T_{\alpha\gamma}{}^{a} \tag{72.7}$$

where

$$T_{\alpha\gamma}{}^{s} = \frac{1}{2}(T_{\alpha\gamma} + T_{\gamma\alpha}) \qquad T_{\alpha\gamma}{}^{a} = \frac{1}{2}(T_{\alpha\gamma} - T_{\gamma\alpha})$$

It is easy to see that the property of symmetry of a tensor is invariant under the transformation of coordinates, i.e., a tensor that is symmetric in one coordinate system will be symmetric in all coordinate systems. The analogous statement for antisymmetric tensors is also correct.

Addition and Subtraction of Tensors. The sum of two tensors $T_{\alpha\gamma}$ and $P_{\alpha\gamma}$ is defined to be a tensor $G_{\alpha\gamma}$ whose components are equal to the sum of the corresponding components of the tensors to be added

$$G_{\alpha\gamma} = T_{\alpha\gamma} + P_{\alpha\gamma} \tag{72.8}$$

The difference of two tensors is defined in a similar manner.

Multiplication of Tensors. The product of a tensor of m^{th} rank $T_{\alpha_1\alpha_2\ldots\alpha_m}$,

and a tensor of the n^{th} rank $P_{\gamma_1\gamma_2\ldots\gamma_n}$ is a tensor of the $(m+n)^{\text{th}}$ rank $R_{\alpha_1\alpha_2\ldots\alpha_m\gamma_1\gamma_2\ldots\gamma_n}$, whose components are equal to the products of the corresponding components of the tensors to be multiplied

$$R_{\alpha_1\alpha_2\ldots\alpha_m\gamma_1\gamma_2\ldots\gamma_n} = T_{\alpha_1\alpha_2\ldots\alpha_m}P_{\gamma_1\gamma_2\ldots\gamma_n} \tag{72.9}$$

Operation of Tensor Contraction. The rank of a tensor may be reduced by two using the operation of "contraction" by summing over any pair of indices. For example, from a fourth-rank tensor $T_{\alpha,\gamma,\mu,\nu}$ we can form the second-rank tensor

$$T_{\alpha\gamma} = \sum_{\lambda=1}^{4} T_{\alpha\gamma\lambda\lambda} \tag{72.10}$$

or

$$T_{\mu\nu} = \sum_{\lambda=1}^{4} T_{\lambda\lambda\mu\nu}$$

This operation is called the *contraction of a tensor*. The result of contracting a second-rank tensor is a scalar.

§73. Tensor Analysis

Four-Dimensional Gradient. Consider some scalar function of four variables $\varphi(x_1, x_2, x_3, x_4)$. The total differential of this function on displacement to another world point infinitesimally close to the initial point equals

$$d\varphi = \frac{\partial\varphi}{\partial x_1}dx_1 + \frac{\partial\varphi}{\partial x_2}dx_2 + \frac{\partial\varphi}{\partial x_3}dx_3 + \frac{\partial\varphi}{\partial x_4}dx_4 \tag{73.1}$$

The increment $d\varphi$ on displacement from one point to another is an invariant, and does not depend on the coordinate system in which it is calculated. The right-hand side of (73.1) is constructed by analogy to the scalar product of an infinitesimal vector dx_1, dx_2, dx_3, dx_4 and of a quantity with components $(\partial\varphi/\partial x_1, \partial\varphi/\partial x_2, \partial\varphi/\partial x_3, \partial\varphi/\partial x_4)$. Since (73.1) is invariant, we conclude that

$$\frac{\partial\varphi}{\partial x_1}, \frac{\partial\varphi}{\partial x_2}, \frac{\partial\varphi}{\partial x_3}, \frac{\partial\varphi}{\partial x_4} \tag{73.2}$$

is a four-dimensional vector, which we call the *four-dimensional gradient* of the function φ. This four-dimensional gradient may be formally represented as the product of the four-dimensional vector operator

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_4} \tag{73.3}$$

and the scalar φ.

The vector nature of the operator (73.3) may be verified directly. We find that

$$\frac{\partial\varphi}{\partial x'_\alpha} = \sum_\gamma \frac{\partial\varphi}{\partial x_\gamma}\frac{\partial x_\gamma}{\partial x'_\alpha} = \sum_\gamma a_{\alpha\gamma}\frac{\partial\varphi}{\partial x_\gamma} \tag{73.3a}$$

using (70.9) to evaluate $\partial x_\gamma/\partial x'_\alpha$. Hence, we may write

$$\frac{\partial}{\partial x'_\alpha} = \sum_\gamma a_{\alpha\gamma}\frac{\partial}{\partial x_\gamma} \tag{73.3b}$$

Thus, it is immediately clear that the operator of the derivatives is transformed according to the vector transformation formulas.

Four-Dimensional Divergence. The scalar product of the four-dimensional vector operator (73.3) and a four-dimensional vector A_1, A_2, A_3, A_4 is an invariant

$$\frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2} + \frac{\partial A_3}{\partial x_3} + \frac{\partial A_4}{\partial x_4} = \text{inv} \tag{73.4}$$

which is called the *four-dimensional divergence* of the vector A_1, A_2, A_3, A_4.

D'Alembert's Operator. Taking the vector A_1, A_2, A_3, A_4 in (73.4) to be the four-dimensional gradient (73.2), we obtain the invariant differential expression

$$\frac{\partial^2\varphi}{\partial x_1^2} + \frac{\partial^2\varphi}{\partial x_2^2} + \frac{\partial^2\varphi}{\partial x_3^2} + \frac{\partial^2\varphi}{\partial x_4^2} = \text{inv} \tag{73.5}$$

The operator

$$\Box = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} + \frac{\partial^2}{\partial x_4^2} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \tag{73.6}$$

is called *D'Alembert's operator.* Its invariance under the Lorentz transformation is due to the vector nature of the operator (73.3), of which it is the square. We have already proved the invariant nature of the latter operator under the Lorentz transformation (see (65.30)) by direct calculation.

Differentiation of Tensors. If we differentiate a tensor with respect to its coordinates, we obtain a tensor which is one higher in rank than the original tensor. The proof of this statement follows from (73.3b). For example, if we differentiate a scalar, we obtain a tensor of the first rank, i.e., a vector, (73.2), which is known as the *gradient.*

Four-Dimensional Curl. By differentiation of A_α, we can form the antisymmetric tensor

$$F_{ij} = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \tag{73.7}$$

which is called the *four-dimensional curl* of A_α. Its space components $(i, j = 1, 2, 3)$ are identical with the components of curl A, where A is the three-dimensional part of the four-dimensional vector A_α.

§74. Tensor Calculus as a Tool of the Theory of Relativity

The principle of relativity requires that the laws of nature should have the same form in all inertial coordinate systems. In other words, the equations which express the laws of nature must be covariant under the transformations which connect the various inertial systems of coordinates. In the special theory of relativity, these transformations are given by the Lorentz transformation, i.e., we require the equations expressing the laws of nature to be covariant under the Lorentz transformation.

It is impossible to tell from the form of an equation whether it is covariant or not. This must be verified by calculations, which are often very cumbersome and laborious. Hence, we need to develop a mathematical tool in which the equations themselves will be of the form which shows immediately whether or not they are relativistically covariant. Such a tool is provided by *tensor calculus*, since tensor equations have the same form in all coordinate systems. For example, suppose that a certain law is expressed in a given coordinate system by the tensor equation

$$T_{\alpha_1, \alpha_2, \ldots, \alpha_n} = 0 \tag{74.1}$$

Multiplying both sides of this equation by the coefficients

$$a_{\alpha'_1, \alpha_1} a_{\alpha'_2, \alpha_2} \cdots a_{\alpha'_n, \alpha_n}$$

and summing over $\alpha_1, \alpha_2, \ldots, \alpha_n$, we obtain

$$\sum_{\alpha_1, \alpha_2, \ldots, \alpha_n} a_{\alpha'_1, \alpha_1} a_{\alpha'_2, \alpha_2} \cdots a_{\alpha'_n, \alpha_n} T_{\alpha_1, \alpha_2, \ldots, \alpha_n} = 0 \tag{74.2}$$

Taking into account the formulas for transforming tensors (72.4), we see that equation (74.2) may be written

$$T'_{\alpha'_1, \alpha'_2, \ldots, \alpha'_n} = 0 \tag{74.3}$$

This is the tensor equation (74.1) transformed into the coordinate system with primes. But (74.3) has the same form as equation (74.1) in the system without primes. Hence, the tensor equation is invariant in form. Thus, by putting an equation in the tensor form, we prove that this equation is covariant.

Thus, for an equation to be covariant, it is necessary and sufficient that it is capable of expression in the tensor form. The problem of deciding

whether this or that equation is covariant is, therefore, reduced to the problem of writing this equation in the tensor form, or, of proving it impossible to do so. Tensor calculus is, therefore, a highly convenient mathematical tool in the theory of relativity.

CHAPTER 16

Relativistic Electrodynamics

§75. Four-Dimensional Potential and Four-Dimensional Current Density

Let us put in four-dimensional form the equations for potentials

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\mu_0\rho\mathbf{v} \tag{75.1}$$

$$\nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -\frac{1}{\epsilon_0}\rho \tag{75.2}$$

and the Lorentz condition associated with them

$$\operatorname{div}\mathbf{A} + \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = 0 \tag{75.3}$$

as well as the equation of continuity

$$\operatorname{div}\rho\mathbf{v} + \frac{\partial\rho}{\partial t} = 0 \tag{75.4}$$

In the coordinates x_1, x_2, x_3, $x_4 = ict$, (75.3) takes the form

$$\frac{\partial\Phi_1}{\partial x_1} + \frac{\partial\Phi_2}{\partial x_2} + \frac{\partial\Phi_3}{\partial x_3} + \frac{\partial\Phi_4}{\partial x_4} = 0 \tag{75.5}$$

where

$$\Phi_1 = A_x \qquad \Phi_2 = A_y \qquad \Phi_3 = A_z \qquad \Phi_4 = \frac{i}{c}\varphi$$

Equation (75.5) is relativistically invariant if Φ_1, Φ_2, Φ_3, Φ_4 form a four-dimensional vector. Under these conditions, the left-hand side of (75.5) is the invariant four-dimensional divergence of this vector. Thus, the

vector potential A and the scalar potential φ form a single four-dimensional Φ_ν with components of the form

$$A_x, A_y, A_z, i\frac{\varphi}{c} \tag{75.6}$$

The vector Φ_ν is called the *four-dimensional potential.*

Equation (75.4) in four-dimensional form becomes

$$\frac{\partial s_1}{\partial x_1} + \frac{\partial s_2}{\partial x_2} + \frac{\partial s_3}{\partial x_3} + \frac{\partial s_4}{\partial x_4} = 0 \tag{75.7}$$

where

$$s_1 = \rho v_x \qquad s_2 = \rho v_y \qquad s_3 = \rho v_z \qquad s_4 = ic\rho$$

If we assume that s_1, s_2, s_3, s_4 form a four-dimensional vector, then equation (75.7) is a relativistically invariant equation expressing the fact that the divergence of the four-dimensional vector s_ν equals zero. The four-dimensional vector s_ν with the components

$$\rho v_x, \rho v_y, \rho v_z, ic\rho \tag{75.8}$$

is called the *four-dimensional current density.*

Using the four-dimensional potential and the four-dimensional current density, we may write equations (75.1) and (75.2) in the form

$$\square \Phi_\nu = -\mu_0 s_\nu \tag{75.9}$$

where $\square$ is d'Alembert's invariant operator (73.6). Hence, (75.9) is, in fact, a tensor equation.

Thus, if we assume that Φ_ν and s_ν are four-dimensional vectors, then equations (75.1) to (75.4) are relativistically invariant equations. But, to verify whether, in fact, they are four-dimensional vectors, we need only to check the numerous consequences of our assumption with experimental fact. This comparison shows that (75.5) and (75.8) are indeed four-dimensional vectors.

Consider a charge of density ρ_0, at rest in the coordinate system denoted by primes. In this system, the components of the four-dimensional current vector are

$$s'_1 = s'_2 = s'_3 = 0 \qquad s'_4 = ic\rho_0 \tag{75.10}$$

In a coordinate system with respect to which the charge moves with a velocity v, the components of the four-dimensional current are, by (71.1)

$$\left.\begin{aligned} s_1 &= \frac{s'_1 - i\beta s'_4}{\sqrt{1-\beta^2}} = \frac{\rho_0 v}{\sqrt{1-\beta^2}} \\ s_2 &= s'_2 = 0 \qquad s_3 = s'_3 = 0 \\ s_4 &= \frac{i\beta s'_1 + s'_4}{\sqrt{1-\beta^2}} = \frac{ic\rho_0}{\sqrt{1-\beta^2}} \end{aligned}\right\} \tag{75.11}$$

Since, by definition, $s_1 = \rho v$ and $s_4 = ic\rho_0$, it follows from these equations that the density of a moving charge is greater than the density of a charge at rest

$$\rho = \frac{\rho_0}{\sqrt{1 - \beta^2}} \tag{75.12}$$

Hence, it follows that the charge dq contained in a given volume is invariant. If a volume dV_0 moves with a velocity v, then, by the Fitzgerald contraction, it becomes

$$dV = dV_0\sqrt{1 - \beta^2} \tag{75.13}$$

This, together with (75.12), gives

$$dq = \rho\, dV = \rho_0\, dV_0 = dq_0 \tag{75.14}$$

i.e., the charge is invariant.

Equations (75.8) for the four-dimensional current density become, taking (75.12) into account

$$\begin{aligned} s_1 &= \frac{\rho_0 v_x}{\sqrt{1 - \beta^2}} & s_2 &= \frac{\rho_0 v_y}{\sqrt{1 - \beta^2}} \\ s_3 &= \frac{\rho_0 v_z}{\sqrt{1 - \beta^2}} & s_4 &= \frac{ic\rho_0}{\sqrt{1 - \beta^2}} \end{aligned} \tag{75.15}$$

Using equations (71.8) for the components of the four-dimensional velocity u_ν, equations (75.15) may be written

$$s_\nu = \rho_0 u_\nu \tag{75.16}$$

In this equation, s_ν and u_ν are four-dimensional vectors, ρ_0 is a scalar. This equation may be considered to be a natural definition of the four-dimensional current density vector.

The fact that charge is independent of velocity is confirmed by many experiments. In particular, if this were not so, no atoms would be neutral. Hence, we use (75.14) as the experimental evidence, and give a strict proof that the quantities s_ν form a four-dimensional vector.

§76. Tensor Form of Maxwell's Equations

In order to formulate Maxwell's theory explicitly in a relativistically invariant form, we must write Maxwell's equations in the tensor form.

Two of Maxwell's equations

$$\begin{aligned} \operatorname{curl} \mathbf{H} &= \rho \mathbf{v} + \frac{\partial \mathbf{D}}{\partial t} \\ \operatorname{div} \mathbf{D} &= \rho \end{aligned} \tag{76.1}$$

may be written as four equations for the components of the vectors

$$\begin{aligned}
0 + \frac{\partial H_z}{\partial x_2} - \frac{\partial H_y}{\partial x_3} - \frac{\partial (icD_x)}{\partial x_4} &= \rho v_x \\
-\frac{\partial H_z}{\partial x_1} + 0 + \frac{\partial H_x}{\partial x_3} - \frac{\partial (icD_y)}{\partial x_4} &= \rho v_y \\
\frac{\partial H_y}{\partial x_1} - \frac{\partial H_x}{\partial x_2} + 0 - \frac{\partial (icD_z)}{\partial x_4} &= \rho v_z \\
\frac{\partial (icD_x)}{\partial x_1} + \frac{\partial (icD_y)}{\partial x_2} + \frac{\partial (icD_z)}{\partial x_3} + 0 &= ic\rho
\end{aligned} \tag{76.2}$$

It follows from (75.8) that the right-hand side of (76.2) contains the components of a four-dimensional vector. Hence, it is clear that the four quantities on the left-hand side of these equations must also form a four-dimensional vector. They all have the form of a four-dimensional divergence (73.4). However, if the four-dimensional divergence of some quantity is a four-dimensional vector, then that quantity itself must be a tensor of the second rank. Equation (76.2) shows that this tensor is of the form

$$F_{\mu\nu} = \begin{bmatrix} 0 & H_z & -H_y & -icD_x \\ -H_z & 0 & H_x & -icD_y \\ H_y & -H_x & 0 & -icD_z \\ icD_x & icD_y & icD_z & 0 \end{bmatrix} \tag{76.3}$$

Using this tensor, (76.2) may be written as the tensor equation

$$\sum_{\nu=1}^{4} \frac{\partial F_{\mu\nu}}{\partial x_\nu} = s_\mu \tag{76.4}$$

where s_μ is defined by (75.8).

Two of Maxwell's vector equations

$$\operatorname{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\operatorname{div} \mathbf{B} = 0$$

may be written in component form

$$\begin{aligned}
\frac{\partial (cB_x)}{\partial x_4} + \frac{\partial (-iE_z)}{\partial x_2} + \frac{\partial (iE_y)}{\partial x_3} &= 0 \\
\frac{\partial (-iE_z)}{\partial x_1} + \frac{\partial (iE_x)}{\partial x_3} + \frac{\partial (-cB_y)}{\partial x_4} &= 0 \\
\frac{\partial (iE_x)}{\partial x_2} + \frac{\partial (cB_z)}{\partial x_4} + \frac{\partial (-iE_y)}{\partial x_1} &= 0 \\
\frac{\partial (cB_z)}{\partial x_3} + \frac{\partial (cB_x)}{\partial x_1} + \frac{\partial (cB_y)}{\partial x_2} &= 0
\end{aligned} \tag{76.5}$$

We shall introduce the tensor

$$H_{\mu\nu} = \begin{bmatrix} 0 & cB_z & -cB_y & -iE_x \\ -cB_z & 0 & cB_x & -iE_y \\ cB_y & -cB_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{bmatrix} \tag{76.6}$$

Using this tensor, we can rewrite (76.5) in the tensor form

$$\frac{\partial H_{\mu\nu}}{\partial x_\lambda} + \frac{\partial H_{\nu\lambda}}{\partial x_\mu} + \frac{\partial H_{\lambda\mu}}{\partial x_\nu} = 0 \tag{76.7}$$

where μ, ν and λ take the values 1, 2, 3, 4; $\mu \neq \nu \neq \lambda$.

There may be more than four possible combinations of the indices in equations (76.7), but there are only four essentially different equations in (76.7). This is evident from the following argument. If there are equal indices, then, since the tensor (76.6) is antisymmetric, the corresponding equation is identically equal to zero. For example, let $\mu = \nu$. Then equation (76.7) is written

$$\frac{\partial H_{\mu\mu}}{\partial x_\lambda} + \frac{\partial H_{\mu\lambda}}{\partial x_\mu} + \frac{\partial H_{\lambda\mu}}{\partial x_\mu} = 0 \tag{76.8}$$

But, this is not an equation but an identity, since

$$H_{\mu\mu} = 0 \qquad H_{\mu\lambda} = -H_{\lambda\mu}$$

Hence, among the equations of (76.7), there are only four equations for which $\mu \neq \nu \neq \lambda$, i.e., in which the indices μ, ν, λ take the following values

$$(\mu, \nu, \lambda) = (2, 3, 4), (3, 4, 1), (4, 1, 2), (1, 2, 3) \tag{76.9}$$

These combinations of the indices in (76.7) give equation (76.5).

We now have to write the equations

$$\mathbf{D} = \epsilon_0 \mathbf{E} \qquad \mathbf{B} = \mu_0 \mathbf{H} \tag{76.10}$$

in invariant form. It is easy to verify immediately that equations (76.10) may be written in the form

$$\left.\begin{aligned} &\frac{1}{c} \sum_{\nu=1}^{4} F_{\mu\nu} u_\nu^{(0)} = \epsilon_0 \sum_{\nu=1}^{4} H_{\mu\nu} u_\nu^{(0)} \\ &\frac{1}{c} (H_{\mu\nu} u_\lambda^{(0)} + H_{\nu\lambda} u_\mu^{(0)} + H_{\lambda\mu} u_\nu^{(0)}) \\ &= \mu_0 (F_{\mu\nu} u_\lambda^{(0)} + F_{\nu\lambda} u_\mu^{(0)} + F_{\lambda\mu} u_\nu^{(0)}) \end{aligned}\right\} \tag{76.11}$$

where $u_\nu^{(0)}$ is a component of the four-dimensional velocity in the frame of reference at rest

$$u_\nu^{(0)} = (0, 0, 0, ic) \tag{76.12}$$

But equation (76.11) is of the tensor form, and, hence, it is also applicable to moving frames of reference. Let the velocity of such a moving frame be v, so that the four-dimensional velocity u_ν has the form

$$u_\nu = \left(\frac{v_x}{\sqrt{1-\beta^2}}, \frac{v_y}{\sqrt{1-\beta^2}}, \frac{v_z}{\sqrt{1-\beta^2}}, \frac{ic}{\sqrt{1-\beta^2}}\right) \tag{76.13}$$

where $\beta = v/c$. Equations (76.11) for a moving frame of reference have the same form, with $u_\nu^{(0)}$ replaced by the four-dimensional velocity u_ν

$$\left.\begin{aligned} \frac{1}{c}\sum_{\nu=1}^{4} F_{\mu\nu}u_\nu &= \epsilon_0 \sum_{\nu=1}^{4} H_{\mu\nu}u_\nu \\ \frac{1}{c}(H_{\mu\nu}u_\lambda + H_{\nu\lambda}u_\mu + H_{\lambda\mu}u_\nu) &= \mu_0(F_{\mu\nu}u_\lambda + F_{\nu\lambda}u_\mu + F_{\lambda\mu}u_\nu) \end{aligned}\right\} \tag{76.14}$$

If these equations are expanded in components and put in the vector form, we obtain

$$\left.\begin{aligned} \mathbf{D} + \frac{1}{c^2}\mathbf{v}\times\mathbf{H} &= \epsilon_0(\mathbf{E} + \mathbf{v}\times\mathbf{B}) \\ \mathbf{B} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E} &= \mu_0(\mathbf{H} - \mathbf{v}\times\mathbf{D}) \end{aligned}\right\} \tag{76.15}$$

When $\mathbf{v} = 0$, these become, as one would expect, equations (76.10) for a frame of reference at rest. We shall discuss the meaning of these equations in greater detail later.

Thus, using the tensors $F_{\mu\nu}$ and $H_{\mu\nu}$ given in (76.3) and (76.6), we may rewrite Maxwell's equations in the tensor form

$$\sum_{\nu=1}^{4} \frac{\partial F_{\mu\nu}}{\partial x_\nu} = s_\mu \tag{76.16}$$

$$\frac{\partial H_{\mu\nu}}{\partial x_\lambda} + \frac{\partial H_{\nu\lambda}}{\partial x_\mu} + \frac{\partial H_{\lambda\mu}}{\partial x_\nu} = 0 \tag{76.17}$$

$$\frac{1}{c}\sum_{\nu=1}^{4} F_{\mu\nu}u_\nu = \epsilon_0 \sum_{\nu=1}^{4} H_{\mu\nu}u_\nu \tag{76.18}$$

$$\frac{1}{c}(H_{\mu\nu}u_\lambda + H_{\nu\lambda}u_\mu + H_{\lambda\mu}u_\nu) = \mu_0(F_{\mu\nu}u_\lambda + F_{\nu\lambda}u_\mu + F_{\lambda\mu}u_\nu) \tag{76.19}$$

The tensors $F_{\mu\nu}$ and $H_{\mu\nu}$ are called the *electromagnetic field tensors.*

§77. Electromagnetic Field Tensors

First of all, we must verify that $F_{\mu\nu}$ and $H_{\mu\nu}$ are tensors. For this purpose, we use the equations which give the electromagnetic field intensity in terms of potentials

$$\mathbf{E} = -\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t} \qquad \mathbf{B} = \operatorname{curl} \mathbf{A} \tag{77.1}$$

Using the four-dimensional potential Φ_ν defined by equation (75.6), we rewrite these equations in the following form

$$\left.\begin{aligned} E_x &= ic\left(\frac{\partial \Phi_4}{\partial x_1} - \frac{\partial \Phi_1}{\partial x_4}\right) & B_x &= \frac{\partial \Phi_3}{\partial x_2} - \frac{\partial \Phi_2}{\partial x_3} \\ E_y &= ic\left(\frac{\partial \Phi_4}{\partial x_2} - \frac{\partial \Phi_2}{\partial x_4}\right) & B_y &= \frac{\partial \Phi_1}{\partial x_3} - \frac{\partial \Phi_3}{\partial x_1} \\ E_z &= ic\left(\frac{\partial \Phi_4}{\partial x_3} - \frac{\partial \Phi_3}{\partial x_4}\right) & B_z &= \frac{\partial \Phi_2}{\partial x_1} - \frac{\partial \Phi_1}{\partial x_2} \end{aligned}\right\} \tag{77.2}$$

Hence, the set of values in (76.6) may be written

$$H_{\mu\nu} = c\left(\frac{\partial \Phi_\nu}{\partial x_\mu} - \frac{\partial \Phi_\mu}{\partial x_\nu}\right) \tag{77.3}$$

But Φ_1, Φ_2, Φ_3, Φ_4 make up a four-dimensional vector. The set of derivatives of the components of a four-dimensional vector, with respect to four-dimensional coordinates, represents a tensor, as was shown in §72. The difference of two tensors is also a tensor. Thus, since Φ_ν is a vector, it follows that $H_{\mu\nu}$ is a tensor.

To show that the set of quantities represented by $F_{\mu\nu}$ is also a tensor, we note that these quantities can be expressed in terms of the tensor $H_{\mu\nu}$. From (76.18), it is clear, since u_μ is a vector and $H_{\mu\nu}$ is a tensor, that $F_{\mu\nu}$ is a tensor.

Thus, we have proved that $F_{\mu\nu}$ and $H_{\mu\nu}$ are tensors. Hence, it follows that in passing from one coordinate system to another, they should be transformed according to the tensor transformation formulas

$$\left.\begin{aligned} F'_{\mu\nu} &= \sum_{\alpha,\gamma} a_{\mu\alpha} a_{\nu\gamma} F_{\alpha\gamma} \\ H'_{\mu\nu} &= \sum_{\alpha,\gamma} a_{\mu\alpha} a_{\nu\gamma} H_{\alpha\gamma} \end{aligned}\right\} \tag{77.4}$$

The coefficients in these formulas are given by the matrix (70.18). Using these values of the coefficients, we may rewrite (77.4) in three-dimensional form

$$\left.\begin{aligned} D'_x &= D_x & H'_x &= H_x \\ D'_y &= \frac{D_y - \frac{v}{c^2} H_z}{\sqrt{1-\beta^2}} & H'_y &= \frac{H_y + vD_z}{\sqrt{1-\beta^2}} \\ D'_z &= \frac{D_z + \frac{v}{c^2} H_y}{\sqrt{1-\beta^2}} & H'_z &= \frac{H_z - vD_y}{\sqrt{1-\beta^2}} \end{aligned}\right\} \tag{77.5}$$

$$\left.\begin{aligned} E'_x &= E_x & B'_x &= B_x \\ E'_y &= \frac{E_y - vB_z}{\sqrt{1-\beta^2}} & B'_y &= \frac{By + \frac{v}{c^2}E_z}{\sqrt{1-\beta^2}} \\ E'_z &= \frac{E_z + vB_y}{\sqrt{1-\beta^2}} & B'_z &= \frac{B_z - \frac{v}{c^2}E_y}{\sqrt{1-\beta^2}} \end{aligned}\right\} \tag{77.6}$$

In these equations, $\beta = v/c$, where v is the velocity of S′ with respect to S. The equations of the inverse transformation are obtained by interchanging the S′ and S terms, and changing the sign of v. It is convenient to write equations (77.5) and (77.6) in the vector form

$$\left.\begin{aligned} \mathbf{D}'_{\parallel} &= \mathbf{D}_{\parallel} & \mathbf{D}'_{\perp} &= \left(\frac{\mathbf{D} + \frac{1}{c^2}\mathbf{v}\times\mathbf{H}}{\sqrt{1-\beta^2}}\right)_{\perp} \\ \mathbf{H}'_{\parallel} &= \mathbf{H}_{\parallel} & \mathbf{H}'_{\perp} &= \left(\frac{\mathbf{H} - \mathbf{v}\times\mathbf{D}}{\sqrt{1-\beta^2}}\right)_{\perp} \end{aligned}\right\} \tag{77.7}$$

$$\left.\begin{aligned} \mathbf{E}'_{\parallel} &= \mathbf{E}_{\parallel} & \mathbf{E}'_{\perp} &= \left(\frac{\mathbf{E} + \mathbf{v}\times\mathbf{B}}{\sqrt{1-\beta^2}}\right)_{\perp} \\ \mathbf{B}'_{\parallel} &= \mathbf{B}_{\parallel} & \mathbf{B}'_{\perp} &= \left(\frac{\mathbf{B} - \frac{1}{c^2}\mathbf{v}\times\mathbf{E}}{\sqrt{1-\beta^2}}\right)_{\perp} \end{aligned}\right\} \tag{77.8}$$

In these expressions, the subscripts "$\parallel$" and "$\perp$" denote that the components are taken, respectively, parallel and perpendicular to the relative motion of the coordinate systems.

These expressions show that (76.15) represents the equations

$$\mathbf{D}' = \epsilon_0\mathbf{E}' \qquad \mathbf{B}' = \mu_0\mathbf{H}'$$

written in the S coordinate system (regarded as being at rest).

Equations (77.5) and (77.6) enable us to find the field vectors in a moving coordinate system if we know them in a system at rest, and vice versa. Thus, we can write as a tensor equation

$$F_{\mu\nu} = \sqrt{\frac{\epsilon_0}{\mu_0}}\,H_{\mu\nu}$$

since, then

$$F'_{\mu\nu} = \sqrt{\frac{\epsilon_0}{\mu_0}}\,H'_{\mu\nu}$$

At low relative velocities of the coordinate systems, when we can ignore terms of the order of $\beta^2 \ll 1$, equations (77.5) and (77.6) assume a much simpler form

$$\mathbf{D}' = \mathbf{D} + \frac{1}{c^2}\mathbf{v} \times \mathbf{H} \qquad \mathbf{H}' = \mathbf{H} - \frac{1}{c^2}\mathbf{v} \times \mathbf{D} \tag{77.9}$$

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B} \qquad \mathbf{B}' = \mathbf{B} - \frac{1}{c^2}\mathbf{v} \times \mathbf{E} \tag{77.10}$$

These equations illustrate very clearly the different roles of the field vectors. They show that the magnetic field vector **H** plays a role analogous to the role of **D** in the electric field, while **B** plays a role in the magnetic field analogous to the role of **E** in the electric field. Hence, if **E** is called the electric field vector, **B** ought to be called the magnetic field vector. However, as we have already pointed out, this name has become firmly attached to **H**, although not entirely correctly. In actual fact, **H** plays the role of the induction. It is impossible now to change the terminology, but we must remember the true significance of the magnetic field vectors.

Invariants of the Field Tensors. Using equations (77.5) to (77.8), we may check by direct calculation that a transition from one coordinate system to another produces no change in the values of the following invariants

$$\left.\begin{aligned} &I_1 = c^2B^2 - E^2 \qquad I'_1 = H^2 - c^2D^2 \\ &I_2 = \mathbf{B}\cdot\mathbf{E} \qquad I'_2 = \mathbf{H}\cdot\mathbf{D} \\ &I_3 = \mathbf{H}\cdot\mathbf{B} - \mathbf{D}\cdot\mathbf{E} \end{aligned}\right\} \tag{77.11}$$

Since these are invariants, we may draw the following conclusions about the behavior of the field vectors under transformation from one coordinate system to another: (1) if in a given coordinate system $c^2B^2 > E^2$ and $\mathbf{B} \perp \mathbf{E}$, then we can choose a coordinate system in which the electric field is zero and the magnetic field nonzero; this cannot be done if **B** is not perpendicular to **E**; (2) if in a given coordinate system $c^2B^2 < E^2$ and $\mathbf{B} \perp \mathbf{E}$, then we can choose a coordinate system such that the magnetic field is zero and the electric field nonzero; this cannot be done if **B** is not perpendicular to **E**; (3) if in a given coordinate system there is either only the electric field or only the magnetic field, then, on transformation to another coordinate system, generally speaking, there will be both the electric field and the magnetic field, and these will be perpendicular to one another; (4) a plane wave for which $cB = E$ and $\mathbf{B} \perp \mathbf{E}$ remains a plane wave in all coordinate systems.

All these statements refer to the behavior of the field vectors at any one point of a four-dimensional space.

§78. Four-Dimensional Force Density

Four-Dimensional Force Density Vector. The Lorentz force density

$$\mathbf{f} = \rho(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{78.1}$$

is written in vector components, taking (76.5) and (75.7) into account, as follows

$$\begin{aligned} f_x &= \frac{1}{c}(\cdots + H_{12}s_2 + H_{13}s_3 + H_{14}s_4) \\ f_y &= \frac{1}{c}(H_{21}s_1 + \cdots + H_{23}s_3 + H_{24}s_4) \\ f_z &= \frac{1}{c}(H_{31}s_1 + H_{32}s_2 + \cdots + H_{34}s_4) \end{aligned} \tag{78.2}$$

The symmetrical structure of these equations suggests the introduction of a four-dimensional force density vector

$$f_\mu = \frac{1}{c}\sum_{\nu=1}^{4} H_{\mu\nu}s_\nu \tag{78.3}$$

The first three components of this force density are identical with the Lorentz force density components (78.2), and the fourth component equals

$$f_4 = \frac{1}{c}(H_{41}s_1 + H_{42}s_2 + H_{43}s_3) = \frac{i}{c}\rho\mathbf{v}\cdot\mathbf{E} \tag{78.4}$$

Using (78.1), f_4 is put in the form

$$f_4 = \frac{i}{c}\mathbf{v}\cdot\mathbf{f} \tag{78.5}$$

Thus, this component is equal, within a factor i/c, to the work carried out by a force of density $\mathbf{f}$ in unit time per unit volume (the work is referred to unit volume because $\mathbf{f}$ is a force density).

Four-Dimensional Minkowski Force Vector. In order to calculate the total force acting on the charge enclosed in a given volume, we must integrate the force density

$$F_\nu = \int f_\nu \, dV \tag{78.6}$$

where dV is a space volume element. However, the quantities F_ν do not form a four-dimensional vector, since the space volume element is not invariant.

To obtain the four-dimensional force vector acting on a point electron, we note that the four-dimensional volume element $dx_1\,dx_2\,dx_3\,dx_4$ is invariant under the Lorentz transformation. To prove this, we write down

the expression for the transformation of volume elements with change of variables

$$dx'_1\,dx'_2\,dx'_3\,dx'_4 = \frac{D(x'_1, x'_2, x'_3, x'_4)}{D(x_1, x_2, x_3, x_4)}\,dx_1\,dx_2\,dx_3\,dx_4 \tag{78.7}$$

where

$$\frac{D(x'_1, x'_2, x'_3, x'_4)}{D(x_1, x_2, x_3, x_4)} = \begin{vmatrix} \frac{\partial x'_1}{\partial x_1} & \frac{\partial x'_2}{\partial x_1} & \frac{\partial x'_3}{\partial x_1} & \frac{\partial x'_4}{\partial x_1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial x'_1}{\partial x_4} & \frac{\partial x'_2}{\partial x_4} & \frac{\partial x'_3}{\partial x_4} & \frac{\partial x'_4}{\partial x_4} \end{vmatrix}$$

is the Jacobian determinant. It is equal to the determinant formed by the coefficients of the transformation (70.18). It can be easily verified, by direct calculation, that this determinant equals 1. Hence, (78.7) takes the form

$$dx'_1\,dx'_2\,dx'_3\,dx'_4 = dx_1\,dx_2\,dx_3\,dx_4 \tag{78.8}$$

which proves the invariance of the four-dimensional volume.

Since a point electron is in a volume element $dV = dx_1\,dx_2\,dx_3$, using the four-dimensional vector of force density f_ν and the invariance of the four-dimensional volume, we can construct a four-dimensional vector of the momentum

$$dP_\nu = f_\nu\,dx_3\,dx_2\,dx\,dt \tag{78.9}$$

Dividing both sides of this equation by the invariant element of proper time of the electron $d\tau$, and integrating with respect to the space coordinates, we obtain the four-dimensional force vector acting on the electron

$$K_\nu = \int \frac{dP_\nu}{d\tau} = \int f_\nu\,dV\,\frac{dt}{d\tau} = \int f_\nu \frac{dV}{\sqrt{1-\beta^2}} \tag{78.10}$$

since $d\tau = dt\sqrt{1-\beta^2}$. The four-dimensional force K_ν is called the *Minkowski force.*

For a point electron, the integration with respect to volume in (78.10) is reduced to an application of the formula

$$\int \rho\,dV = e \tag{78.11}$$

Hence, the first three components of the Minkowski force are equal to the three components of the Lorentz force

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{78.12}$$

divided by $\sqrt{1-\beta^2}$, while the fourth component, on the basis of (78.4), equals

$$K_4 = \frac{i}{c}\frac{e}{\sqrt{1-\beta^2}}\mathbf{v}\cdot\mathbf{E} \tag{78.13}$$

Hence, the four components of the Minkowski force may be written

$$K = \left(\frac{\mathbf{F}}{\sqrt{1-\beta^2}},\ \frac{i}{c}\frac{e(\mathbf{v}\cdot\mathbf{E})}{\sqrt{1-\beta^2}}\right) \tag{78.14}$$

where the Lorentz force $\mathbf{F}$ is defined by (78.12).

§79. Electromagnetic Field Energy Momentum Tensor

In order to complete the presentation of the electromagnetic field quantities and equations for empty space in a relativistically invariant form, it is necessary to find a relativistically invariant representation of the density of the energy and momentum of the electromagnetic field. The work done by the electromagnetic field, according to the law of conservation of energy and momentum, is accompanied by corresponding changes in the energy and momentum of the electromagnetic field. Hence, the four-dimensional force vector f_ν must be related to the density of the energy and momentum of the electromagnetic field. Therefore, we shall express this vector in terms of quantities describing the electromagnetic field, i.e., in terms of the field tensors $F_{\mu\nu}$ and $H_{\mu\nu}$.

Substituting the value of s_ν from Maxwell's equations (76.16) in the expression for the force density f_μ (78.3), we obtain

$$f_\mu = \frac{1}{c}\sum_{\alpha,\gamma} H_{\mu\alpha}\frac{\partial F_{\alpha\gamma}}{\partial x_\gamma} \tag{79.1}$$

Using the self-evident equation

$$H_{\mu\alpha}\frac{\partial F_{\alpha\gamma}}{\partial x_\gamma} = \frac{\partial(H_{\mu\alpha}F_{\alpha\gamma})}{\partial x_\gamma} - \frac{\partial H_{\mu\alpha}}{\partial x_\gamma}F_{\alpha\gamma} \tag{79.2}$$

and, since $H_{\mu\nu}$ and $F_{\mu\nu}$ are antisymmetric

$$\frac{\partial H_{\mu\alpha}}{\partial x_\gamma}F_{\alpha\gamma} = \frac{1}{2}\left(\frac{\partial H_{\mu\alpha}}{\partial x_\gamma}F_{\alpha\gamma} + \frac{\partial H_{\alpha\mu}}{\partial x_\gamma}F_{\gamma\alpha}\right) \tag{79.3}$$

From (79.3), it follows that

$$\begin{aligned}\sum_{\alpha,\gamma}\frac{\partial H_{\mu\alpha}}{\partial x_\gamma}F_{\alpha\gamma} &= \frac{1}{2}\sum_{\alpha,\gamma}\left(\frac{\partial H_{\mu\alpha}}{\partial x_\gamma}F_{\alpha\gamma} + \frac{\partial H_{\alpha\mu}}{\partial x_\gamma}F_{\gamma\alpha}\right)\\ &= \frac{1}{2}\sum_{\alpha,\gamma}F_{\alpha\gamma}\left(\frac{\partial H_{\mu\alpha}}{\partial x_\gamma} + \frac{\partial H_{\gamma\mu}}{\partial x_\alpha}\right)\end{aligned} \tag{79.4}$$

with the replacement of the summation indices $\alpha \to \gamma$ and $\gamma \to \alpha$ in the second term, i.e.

$$\sum_{\alpha,\gamma} \frac{\partial H_{\alpha\mu}}{\partial x_\gamma} F_{\gamma\alpha} = \sum_{\alpha,\gamma} \frac{\partial H_{\gamma\mu}}{\partial x_\alpha} F_{\alpha\gamma} \tag{79.5}$$

The expression in the last set of parentheses in (79.4) equals $-\partial H_{\alpha\gamma}/\partial x_\mu$ on the basis of Maxwell's equation (76.7). Consequently

$$\sum_{\alpha,\gamma} \frac{\partial H_{\mu\alpha}}{\partial x_\gamma} F_{\alpha\gamma} = -\frac{1}{2} \sum_{\alpha,\gamma} F_{\alpha\gamma} \frac{\partial H_{\alpha\gamma}}{\partial x_\mu}$$

To evaluate this expression, we use (77.4). Then we obtain

$$\sum_{\alpha\gamma} F_{\alpha\gamma} \frac{\partial H_{\alpha\gamma}}{\partial x_\mu} = \sum_{\alpha,\gamma} \sqrt{\frac{\epsilon_0}{\mu_0}} H_{\alpha\gamma} \frac{\partial H_{\alpha\gamma}}{\partial x_\mu} = \frac{1}{2} \frac{\partial}{\partial x_\mu} \sum_{\alpha,\gamma} \sqrt{\frac{\epsilon_0}{\mu_0}} H_{\alpha\gamma} H_{\alpha\gamma}$$

$$= \frac{1}{2} \frac{\partial}{\partial x_\mu} \sum_{\alpha,\gamma} F_{\alpha\gamma} H_{\alpha\gamma} \tag{79.6}$$

Hence, taking (79.2), (79.4), and (79.6) into account, we may write (79.1) in the form

$$f_\mu = \frac{1}{c} \sum_{\gamma=1}^{4} \frac{\partial}{\partial x_\gamma} \left(\sum_{\alpha=1}^{4} H_{\mu\alpha} F_{\alpha\gamma} \right) + \frac{1}{4c} \frac{\partial}{\partial x_\mu} \sum_{\alpha,\gamma} (F_{\alpha\gamma} H_{\alpha\gamma}) \tag{79.7}$$

We introduce the notation

$$T_{\mu\nu} = \frac{1}{c} \sum_{\alpha=1}^{4} H_{\mu\alpha} F_{\alpha\nu} + \delta_{\mu\nu} \frac{1}{4c} \sum_{\alpha=1}^{4} \sum_{\gamma=1}^{4} (F_{\alpha\gamma} H_{\alpha\gamma}) \tag{79.8}$$

where $\delta_{\mu\nu}$ is Kronecker's delta, defined by

$$\delta_{\mu\nu} = \begin{cases} 0 & \text{for} \quad \mu \neq \nu \\ 1 & \text{for} \quad \mu = \nu \end{cases}$$

Then the expression for the force density f_μ may be put in the form

$$f_\mu = \sum_{\nu=1}^{4} \frac{\partial T_{\mu\nu}}{\partial x_\nu} \tag{79.9}$$

i.e., in the form of the four-dimensional divergence of a tensor $T_{\mu\nu}$. $T_{\mu\nu}$ is called the *energy-momentum tensor of the electromagnetic field*. It is evident from (79.8) that this tensor is symmetric. The expression for the components of this tensor in terms of the electromagnetic field intensity may be obtained by substituting the expressions for $F_{\mu\nu}$ and $H_{\mu\nu}$ from (76.3) and (76.6) in (79.8).

The following expression is obtained for the constant term of (79.8)

$$\frac{1}{4c} \sum_{\alpha,\gamma} (F_{\alpha\gamma} H_{\alpha\gamma}) = \frac{1}{2} \{(\mathbf{H} \cdot \mathbf{B}) - (\mathbf{D} \cdot \mathbf{E})\} \tag{79.10}$$

The remaining terms of the various components of the tensor (79.8) are calculated in a similar manner using (76.3) and (76.6). We shall use the notation

$$\left.\begin{aligned}
T_{xx} &= D_xE_x + H_xB_x - \frac{1}{2}(\mathbf{D}\cdot\mathbf{E} + \mathbf{H}\cdot\mathbf{B}) \\
T_{xy} &= T_{yx} = D_xE_y + H_xB_y = E_xD_y + B_xH_y \\
T_{yy} &= D_yE_y + H_yB_y - \frac{1}{2}(\mathbf{D}\cdot\mathbf{E} + \mathbf{H}\cdot\mathbf{B}) \\
T_{yz} &= T_{zy} = D_yE_z + H_yB_z = D_zE_y + B_yH_z \\
T_{zz} &= D_zE_z + H_zB_z - \frac{1}{2}(\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B}) \\
T_{zx} &= T_{xz} = E_zD_x + B_zH_x = D_zE_x + B_xH_z
\end{aligned}\right\} \tag{79.11}$$

Using the notation of (79.11), and the fact that

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \qquad \mathbf{g} = \frac{1}{c^2}\mathbf{E} \times \mathbf{H}$$

$$u = \frac{1}{2}(\mathbf{D}\cdot\mathbf{E} + \mathbf{H}\cdot\mathbf{B})$$

we may write the energy-momentum tensor in the form

$$T_{\mu\nu} = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} & -icg_x \\ T_{yx} & T_{yy} & T_{yz} & -icg_y \\ T_{zx} & T_{zy} & T_{zz} & -icg_z \\ -\frac{i}{c}P_x & -\frac{i}{c}P_y & -\frac{i}{c}P_z & u \end{bmatrix} \tag{79.12}$$

We shall now clarify the meaning of the various components of $T_{\mu\nu}$, and the significance of the equations (79.9).

Let us integrate the fourth equation of (79.9) over the three-dimensional volume $dV = dx_1\,dx_2\,dx_3$

$$\int f_4\,dV = \int \left(\frac{\partial T_{41}}{\partial x} + \frac{\partial T_{42}}{\partial y} + \frac{\partial T_{43}}{\partial z}\right) dV + \frac{1}{ic}\frac{d}{dt}\int u\,dV \tag{79.13}$$

Using the expression (78.5) for f_4, and expressing the components of T_{4i} in terms of (79.12), (79.13) may be rewritten

$$\begin{aligned}
\int_V \mathbf{v}\cdot\mathbf{f}\,dV &= -\int_V \operatorname{div}\mathbf{P}\,dV - \frac{d}{dt}\int u\,dV \\
&= -\int_S \mathbf{P}\cdot d\mathbf{S} - \frac{d}{dt}\int u\,dV
\end{aligned} \tag{79.14}$$

where the volume integral is transformed into a surface integral by means of Gauss' theorem. Equation (79.14) is identical with equation (8.11), which expresses the law of conservation of energy of the electromagnetic field.

Thus, equation (79.9) with $\mu = 4$ expresses the law of conservation of energy of the electromagnetic field. This is the justification for the insertion, in the last row of (79.12), of the quantities P_x, P_y, P_z, which describe the energy flow, and not g_x, g_y, g_z, which describe the momentum, although

$$-\frac{i}{c}\mathbf{P} = -ic\mathbf{g} \tag{79.15}$$

and formally we could replace the three terms in the last column of (79.12) with the three terms in the last row.

We shall now consider one of the first three equations in (79.9), e.g., the first, with $\mu = 1$. Integrating, as before, over the three-dimensional volume

$$\int f_1\,dV = \int \left(\frac{\partial T_{11}}{\partial x} + \frac{\partial T_{12}}{\partial y} + \frac{\partial T_{13}}{\partial z}\right) dV + \frac{1}{ic}\frac{d}{dt}\int T_{14}\,dV \tag{79.16}$$

Since $f_1 = f_x$ is the x component of the force density, then the left-hand side contains the force acting on the charge in a given volume. The last term of (79.16)

$$\frac{1}{ic}\frac{d}{dt}\int T_{14}\,dV = -\frac{d}{dt}\int g_x\,dV$$

gives the change in the x component of the electromagnetic field momentum in the volume under consideration. The terms T_{1i}, occurring in the first integral of (79.16), may be considered as various components of the vector F

$$\mathsf{F}_x = T_{11} \qquad \mathsf{F}_y = T_{12} \qquad \mathsf{F}_z = T_{13} \tag{79.17}$$

Then, by Gauss' formula, this integral may be transformed into a surface integral

$$\int_V \left(\frac{\partial T_{11}}{\partial x} + \frac{\partial T_{12}}{\partial y} + \frac{\partial T_{13}}{\partial z}\right) dV = \int_V \operatorname{div}\mathsf{F}\,dV = \int_S \mathsf{F}\cdot d\mathbf{S} \tag{79.18}$$

The left-hand side of equation (79.16) contains a force, and, therefore, the integral (79.18) must have the dimensions of force. Therefore, we may conclude that this is the x component of the surface force acting on the surface of the volume under consideration.

According to the laws of classical mechanics, the force

$$\int f_1\,dV \tag{79.19}$$

produces a change in the momentum of material bodies. Denoting this momentum in the volume under consideration by $\mathbf{G}^M$, we have

$$\frac{dG_x^M}{dt} = \int f_1\,dV \tag{79.20}$$

Hence, equation (79.16) may be written

$$\frac{d}{dt}\left(G_x^M + \int_V g_x\,dV\right) = \int_S \mathbf{F}\cdot d\mathbf{S} \tag{79.21}$$

This equation is identical with equation (41.9), which has been obtained by direct calculation of the forces acting on the charges within a given volume.

Equations (79.9) for $\mu = 2, 3$ are similar in form. These equations express the law of conservation of momentum, since they relate the total force acting on a volume (given in the right-hand side of equation (79.21)), to the change of the momentum of bodies and of the electromagnetic field within that volume.

Thus, consideration of the physical meaning of the fourth equation of (79.9) leads to the interpretation that P_x, P_y, P_z are components of the density of the electromagnetic energy flux and u is the electromagnetic energy density. Consideration of the physical meaning of the first three equations of (79.9) leads to the interpretation that g_x, g_y, g_z are components of the density of the electromagnetic momentum. The quantities T_{xx}, T_{xy}, etc., describe the surface forces applied to the surface of the volume under consideration. Therefore, the three-dimensional tensor

$$T = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} \tag{79.22}$$

is called the *stress tensor*. The forces produced by this tensor are evaluated by equations of the form (79.18).

Equation (78.19) shows that the energy-momentum tensor is symmetric. From this symmetry, irrespective of the meaning of $\mathbf{P}$ and $\mathbf{g}$, follows the fundamental relationship between the density of the energy flux and the density of the electromagnetic momentum

$$\mathbf{g} = \frac{\mathbf{P}}{c^2} \tag{79.23}$$

§80. Doppler Effect

Invariance of Plane Wave. As we have already observed, the concept of a plane wave is invariant under the Lorentz transformation: a plane wave remains a plane wave in all coordinate systems. In the case of a plane

wave, the vectors $\mathbf{E}$ and $\mathbf{H}$ are defined by the amplitudes $\mathbf{E}_0$ and $\mathbf{H}_0$ and the phase $\varphi = \omega t - \mathbf{k}\cdot\mathbf{r}$

$$\mathbf{E} = \mathbf{E}_0 \sin(\omega t - \mathbf{k}\cdot\mathbf{r}) \qquad \mathbf{H} = \mathbf{H}_0 \sin(\omega t - \mathbf{k}\cdot\mathbf{r}) \tag{80.1}$$

Invariance of Phase. The phase of a wave is invariant under the Lorentz transformation. This may be shown in the following manner. The statement that at some space-time point the vectors $\mathbf{E}$ and $\mathbf{H}$ of the wave are zero is true in all coordinate systems. But, this means that the phase of the wave has the same value in all coordinate systems, and that this value is an integral multiple of π. This proves the invariance of the phase. The invariance of the phase also follows from the transformation formulas for the field vectors. If we write down these well-known formulas, and substitute in them the expressions for the field vectors of a plane wave (80.1), we may at once conclude that the phases must be equal for these formulas to be true at all instants of time.

Four-Dimensional Wave Vector. The phase may be written in the form

$$-\varphi = \mathbf{k}\cdot\mathbf{r} - \omega t = k_1x_1 + k_2x_2 + k_3x_3 + k_4x_4 \tag{80.2}$$

The product (80.2) is invariant. It has the form of the scalar product of the four-dimensional vector x_1, x_2, x_3, x_4, and the set of quantities k_1, k_2, k_3, k_4

$$k_1 = k_x \qquad k_2 = k_y \qquad k_3 = k_z \qquad k_4 = i\frac{\omega}{c} \tag{80.3}$$

Since (80.2) is invariant and the set x_1, x_2, x_3, x_4 constitutes a four dimensional vector, it follows that the set k_1, k_2, k_3, k_4 also constitutes a four-dimensional vector called the *wave vector* k.

Expressions for Transforming Frequency and Direction. If $\mathbf{n}$ denotes a unit vector in the direction of propagation of a plane wave, then the wave vector $\mathbf{k}$ is of the form

$$\mathbf{k} = \frac{\omega}{c}\mathbf{n} \tag{80.4}$$

Since k_1, k_2, k_3, k_4 is a four-dimensional vector, then, using the transformation formulas of vectors, we obtain the following expressions for transforming the frequency and direction of a plane wave

$$\left.\begin{aligned} \omega n_x &= \omega' \frac{\beta + n'_x}{\sqrt{1-\beta^2}} \\ \omega n_y &= \omega' n'_y \\ \omega n_z &= \omega' n'_z \\ \omega &= \omega' \frac{1 + \beta n'_x}{\sqrt{1-\beta^2}} \end{aligned}\right\} \tag{80.5}$$

Doppler Effect. The *Doppler effect* is the name given to the change in the frequency of light due to the motion of its source. The fourth equation in (80.5) describes the Doppler effect. It differs from the classical expression of the Doppler effect by the denominator $\sqrt{1-\beta^2}$, which takes into account the effect of time dilatation. Hence, the experimental verification of the relativistic Doppler effect must confirm the effect of time dilatation. Such an experiment was performed by Ives in 1938.

From the last equation of (80.5), using the principle of relativity for the first equation of (80.5), we obtain

$$\omega = \omega' \frac{\sqrt{1-\beta^2}}{1-\beta n_x} \tag{80.6}$$

which describes the Doppler effect: n_x is the cosine of the angle between the direction of propagation of the light and the x axis. The source is assumed to be moving with a velocity v ($\beta = v/c$).

Longitudinal Doppler Effect. If the direction of the light ray and the direction of the motion of the source are the same (in this case, the positive x direction), then we obtain the *longitudinal Doppler effect,* familiar from classical optics. In this case, $n_x = n'_x = \pm 1$, and (80.6) becomes, for $\beta \ll 1$

$$\omega = \omega' \frac{\sqrt{1-\beta^2}}{1 \mp \beta} \approx \omega'(1 \pm \beta) \tag{80.7}$$

where only the term linear in β is retained. This equation is the same as the classical expression for the Doppler effect, which has been confirmed experimentally.

Transverse Doppler Effect. When the source of light moves in a direction perpendicular to that of the observer ($n_x = 0$), we have the *transverse Doppler effect* ($\beta \ll 1$)

$$\omega = \omega'\sqrt{1-\beta^2} \approx \omega'\left(1 - \frac{1}{2}\beta^2\right) \tag{80.8}$$

This is an effect of the second order of smallness in $\beta = v/c$. It is entirely due to the presence of the term $\sqrt{1-\beta^2}$ in equations (80.5), and is a purely relativistic effect associated with the time dilatation of the moving source. Experimental confirmation of this effect was obtained by Ives.

Ives' Experiment (1938). Ives used the radiation of hydrogen atoms moving at velocities up to 1.8×10^8 cm/sec ($\beta \approx 6 \times 10^{-3}$). In the observation of the Doppler effect, we deal simultaneously with the effect proportional to β and the effect proportional to β^2. It is clear that it would be quite difficult to observe the effect of (80.8); it would be necessary to find the angle between the direction of motion of the particle and the

normal to this motion, which is also the direction of observation (with accuracy to an angle $\alpha \ll 6 \times 10^{-3}$). Therefore, Ives observed the influence of the relativistic term in the Doppler effect in a different way. He observed the total Doppler effect in the direction of motion and in the opposite direction. Two lines, whose frequencies are

$$\omega_1 = \omega' \frac{1 + \beta}{\sqrt{1 - \beta^2}} \qquad \omega_2 = \omega' \frac{1 - \beta}{\sqrt{1 - \beta^2}} \tag{80.9}$$

as well as the line of the atom at rest, $\omega_0 = \omega'$, were photographed on the same plate. The mean frequency of the displaced lines (80.9) was

$$\langle \omega \rangle = \omega' \frac{1}{\sqrt{1 - \beta^2}} \tag{80.10}$$

The position of this mean frequency $\langle \omega \rangle$ was measured with respect to the undisplaced frequency ω_0. This confirmed the relationship (80.10), i.e., the influence of the term $\sqrt{1 - \beta^2}$ in the expressions for the Doppler effect. This means that the Ives' experiment confirmed both the transverse Doppler effect (80.8) and the time dilatation of moving clocks.

§81. Plane Waves

Transformation of the Amplitude and Frequency of Plane Waves. Let us consider the case when the wave vector (80.4) of a plane wave lies in the xy plane (Fig. 78). We then have the following expressions for the intensities in the plane wave

$$\begin{aligned} E_x &= -an_y e^{i\varphi} \\ E_y &= an_x e^{i\varphi} \\ H_z &= ae^{i\varphi} \end{aligned} \tag{81.1}$$

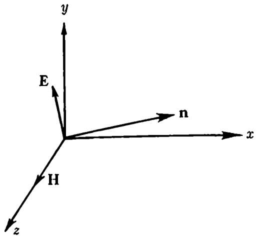


Fig. 78

where φ is the phase of the wave, which is invariant under the Lorentz transformation. Using equations (77.5) and (77.6), we obtain the following expressions for the transformation of the amplitude and the direction of the normal of a plane wave

$$\begin{aligned} a &= a' \frac{1 + \beta n'_x}{\sqrt{1 - \beta^2}} \\ n_x a &= a' \frac{n'_x + \beta}{\sqrt{1 - \beta^2}} \\ n_y a &= n'_y a' \end{aligned} \tag{81.2}$$

Dividing the second and third equations by the first, we obtain the following expressions for the transformation of the components of the vector normal to the plane wave

$$n_x = \frac{n'_x + \beta}{1 + \beta n'_x} \qquad n_y = \frac{n'_y\sqrt{1 - \beta^2}}{1 + \beta n'_x} \tag{81.3}$$

These expressions are identical with the corresponding expressions obtained in a similar manner from (80.5). On the other hand, dividing the first equation of (81.2) by the last equation of (80.5), we obtain the following relationship between the amplitudes and frequencies of plane wave

$$\frac{a}{\omega} = \frac{a'}{\omega'} = \text{inv} \tag{81.4}$$

which is invariant under the Lorentz transformation.

Energy of a Plane Wave. From (8.12), it follows that the energy of a wave train contained in a volume V is equal to

$$W = \frac{\epsilon}{2} a^2 V \tag{81.5}$$

The volume occupied by the wave train moves with the velocity of light; hence, it is impossible to relate this volume to a coordinate system, or to speak of its rest volumes. We introduce an auxiliary volume V, which moves with a velocity u' in a coordinate system S′ and with a velocity u in a coordinate system S. From the Fitzgerald contraction of the volume, we have

$$V' = V_0\sqrt{1 - \frac{u'^2}{c^2}} \qquad V = V_0\sqrt{1 - \frac{u^2}{c^2}} \tag{81.6}$$

Hence, it follows that

$$\frac{V'}{V} = \frac{\sqrt{1 - \frac{u'^2}{c^2}}}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{81.7}$$

Hence, using equation (71.11c), we find

$$V' = V\frac{\sqrt{1-\beta^2}}{1-\dfrac{vu_x}{c^2}} \tag{81.8}$$

We shall substitute u_x in (81.8), using

$$u_x = un_x \tag{81.9}$$

where n_x is the projection of the unit vector in the direction of **u** onto the x axis. Then, we allow u to tend to c, and we obtain

$$V' = V\frac{\sqrt{1-\beta^2}}{1-\beta n_x} \tag{81.10}$$

The first expression in (81.2) may be rewritten, in accordance with the principle of relativity

$$a' = a\frac{1-\beta n_x}{\sqrt{1-\beta^2}} \tag{81.11}$$

Comparing equations (81.10), (81.11), and (80.6), we obtain

$$\frac{W'}{\omega'} = \frac{W}{\omega} \qquad V'\omega' = V\omega \tag{81.12}$$

Hence, the energy of a train of plane waves is transformed like the frequency; in other words, the energy is directly proportional to the frequency. This is used in quantum theory.

Momentum of a Plane Wave. The momentum of a train of plane waves equals

$$\mathbf{G} = \frac{1}{c^2}\int \mathbf{E}\times\mathbf{H}\,dV = \frac{\mathbf{n}}{c}\int \epsilon_0 E^2\,dV = \mathbf{n}\frac{W}{c} \tag{81.13}$$

where **n** is the unit vector in the direction of propagation. In deducing (81.13), we take into account the following relationships

$$\frac{1}{c^2}EH = \frac{1}{c^2}E\sqrt{\frac{\epsilon_0}{\mu_0}}E = \frac{1}{c}\epsilon_0 E^2$$

Using (81.12), we may rewrite (81.13) in the form

$$\mathbf{G} = \mathbf{n}\omega d \qquad W = \omega c d \tag{81.14}$$

where, according to (81.12), $d = W/\omega c$ is a constant. According to (80.3)

$$\omega n_x,\ \omega n_y,\ \omega n_z,\ -i\omega \tag{81.15}$$

form a four-dimensional vector. Therefore, it follows from (81.14) that the quantities

$$G_x, G_y, G_z, i\frac{W}{c} \tag{81.16}$$

which describe the total momentum and total energy of a train of plane waves also form a four-dimensional vector.

Reflection from a Moving Mirror. Consider a plane mirror, perpendicular to the x axis, moving in the positive x direction with a velocity v. A plane light wave, of frequency ω'_0, strikes the moving mirror. We shall assume that the vector normal to the plane wave lies in the xy plane, and makes an angle Θ_0 with the x axis (Fig. 79).

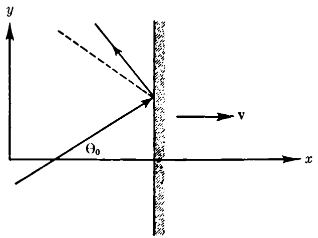


Fig. 79

Applying (80.5), it is easy to obtain analogous expressions for the transformation from a coordinate system at rest to a moving system S′, in which the mirror is at rest

$$\begin{aligned} \omega' &= \omega_0 \frac{1 - \beta n_x}{\sqrt{1-\beta^2}} = \omega_0 \frac{1 - \beta\cos\Theta_0}{\sqrt{1-\beta^2}} \\ n'_x &= \frac{n_x - \beta}{1 - \beta n_x} = \frac{\cos\Theta_0 - \beta}{1 - \beta\cos\Theta_0} \\ n'_y &= \frac{n_y\sqrt{1-\beta^2}}{1 - \beta n_x} = \frac{\sin\Theta_0\sqrt{1-\beta^2}}{1 - \beta\cos\Theta_0} \end{aligned} \tag{81.17}$$

Since the mirror is at rest in S′, the ordinary laws of reflection apply. The frequency is unchanged by reflection, and the angle of incidence equals the angle of reflection. Hence, from (81.17), we immediately obtain for the reflected wave

$$
\begin{aligned}
\omega'_{\mathrm{ref}} &= \omega_0 \frac{1 - \beta \cos \Theta_0}{\sqrt{1 - \beta^2}} \\
n'_{x\,\mathrm{ref}} &= -\frac{\cos \Theta_0 - \beta}{1 - \beta \cos \Theta_0} \\
n'_{y\,\mathrm{ref}} &= \frac{\sin \Theta_0 - \beta}{1 - \beta \cos \Theta}
\end{aligned}
\tag{81.18}
$$

Now, using (80.5), we may return to the S coordinate system, considered to be at rest

$$
\begin{aligned}
\omega_{\mathrm{ref}} &= \omega'_{\mathrm{ref}} \frac{1 + \beta n'_{x\,\mathrm{ref}}}{\sqrt{1 - \beta^2}} = \omega_0 \frac{1 - 2\beta \cos \Theta_0 + \beta^2}{1 - \beta^2} \\
\cos \Theta_{\mathrm{ref}} &= n_{x\,\mathrm{ref}} = \frac{n'_{x\,\mathrm{ref}} + \beta}{1 + \beta n'_{x\,\mathrm{ref}}} = -\frac{\cos \Theta_0 - 2\beta + \beta^2 \cos \Theta_0}{1 - 2\beta \cos \Theta_0 + \beta^2} \\
\sin \Theta_{\mathrm{ref}} &= n_{y\,\mathrm{ref}} = \frac{n'_{y\,\mathrm{ref}}\sqrt{1 - \beta^2}}{1 + \beta n'_{x\,\mathrm{ref}}} = \frac{\sin \Theta_0 (1 - \beta^2)}{1 - 2\beta \cos \Theta_0 + \beta^2}
\end{aligned}
\tag{81.19}
$$

In the nonrelativistic case, when $\beta \ll 1$, these equations may be written, with accuracy to the order of β^2

$$
\begin{aligned}
\omega_{\mathrm{ref}} &\approx \omega_0 (1 - 2\beta \cos \Theta_0) \\
\cos \Theta_{\mathrm{ref}} &\approx -\cos \Theta_0 + 2\beta \sin^2 \Theta_0 \\
\sin \Theta_{\mathrm{ref}} &\approx \sin \Theta_0 \, (1 + 2\beta \cos \Theta_0)
\end{aligned}
$$

In Fig. 79, the broken line represents the ray reflected from a mirror at rest, and the continuous line represents the ray reflected from a moving mirror, in the coordinate system in which the mirror is moving.

§82. Field of an Arbitrarily Moving Electron

Lienard-Wiechert Potentials. Let the coordinates of the electron be $y_1, y_2, y_3, y_4 = ict'$, and the coordinates of the point at which the potential due to the moving electron is evaluated, be $x_1, x_2, x_3, x_4 = ict$. We shall assume that the electromagnetic perturbation caused by the electron is propagated with the velocity of light. This means that

$$(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 = c^2 (t - t')^2 \tag{82.1}$$

If we introduce a four-dimensional vector with the components

$$R_i = x_i - y_i \tag{82.2}$$

then (82.1) may be written in the following relativistically invariant form

$$R^2 \equiv \sum_{i=1}^{4} R_i^2 = 0 \tag{82.3}$$

We now have to find the four-dimensional potential set up by the arbitrarily moving electron. In the coordinate system in which the electron is at rest at a given instant of time, the four-dimensional potential which it sets up is known to be

$$\Phi_1^{(0)} = \Phi_2^{(0)} = \Phi_3^{(0)} = 0 \qquad \Phi_4^{(0)} = i\frac{\varphi^{(0)}}{c} = \frac{i}{c}\frac{1}{4\pi\epsilon_0}\frac{e}{r} \tag{82.4}$$

since, in this case, the electromagnetic field reduces to the Coulomb field of a point charge. In order to obtain expressions for the four-dimensional potential in an arbitrary coordinate system, i.e., a coordinate system in which the electron moves in an arbitrary manner, we must rewrite (82.4) in the form of a single vector equation. We note that in the coordinate system in which, at a given instant, the electron is at rest, the components of the four-dimensional velocity of the electron $u_i = dy_i/d\tau$ are

$$u_1^{(0)} = u_2^{(0)} = u_3^{(0)} \qquad u_4 = ic \tag{82.5}$$

Hence, equation (82.4) may be written in the form of the vector relationships

$$\Phi_\nu = \frac{1}{4\pi\epsilon_0}\frac{eu_\nu}{r_0c^2} \tag{82.6}$$

where r_0 is an invariant, equal to the three-dimensional distance from the electron to the point under discussion in the coordinate system in which the electron is at rest. This invariant may be written

$$r_0 = -\frac{1}{c}\sum_i R_iu_i \tag{82.7}$$

The right-hand side of this equation contains the scalar product of two four-dimensional vectors, which is invariant. It is most convenient to evaluate this product in the coordinate system in which the electron is at rest. But in this system equation (82.7) is obviously true, and hence, it is proved for all coordinate systems. Therefore, using (82.7), (82.6) may be put in the form

$$\Phi_\nu = -\frac{1}{4\pi\epsilon_0 c}\frac{eu_\nu}{\sum_i R_iu_i} \tag{82.8}$$

with the additional condition (82.3). Equation (82.8) is completely relativistically invariant. Rewriting it in the coordinate system in which the electron is moving, we obtain the expression for the potentials set up by the moving electron.

Taking into account the fact that

$$(u_1, u_2, u_3) = \frac{(v_1, v_2, v_3)}{\sqrt{1-\beta^2}} \qquad u_4 = \frac{ic}{\sqrt{1-\beta^2}} \tag{82.9}$$

where $\mathbf{v}$ is the three-dimensional velocity of the electron, and using the equation

$$\sum_{i=1}^{4} R_i u_i = \frac{\mathbf{r}\cdot\mathbf{v}}{\sqrt{1-\beta^2}} - \frac{rc}{\sqrt{1-\beta^2}} \tag{82.10}$$

we can express the expression (82.8) for the four-dimensional potential in the form of the following equations for the vector and scalar potentials

$$\mathbf{A} = \frac{1}{4\pi\epsilon_0}\frac{1}{c^2}\left[\frac{e\mathbf{v}}{r - \dfrac{\mathbf{r}\cdot\mathbf{v}}{c}}\right]_{t-\frac{r}{c}} = \frac{\mu_0}{4\pi}\left[\frac{e\mathbf{v}}{r - \dfrac{\mathbf{r}\cdot\mathbf{v}}{c}}\right]_{t-\frac{r}{c}} \tag{82.11}$$

$$\varphi = \frac{1}{4\pi\epsilon_0}\frac{e}{\left[r - \dfrac{\mathbf{r}\cdot\mathbf{v}}{c}\right]_{t-\frac{r}{c}}} \tag{82.12}$$

In these equations, the potentials A and φ are evaluated at some point of space at time t. The velocity $\mathbf{v}$ of the electron and the radius vector $\mathbf{r}$, drawn from the position of the electron to the point where the potential is to be evaluated, must be taken not at time t, but at $t - r/c$, because of the finite rate of propagation of electromagnetic effects, given by (82.3). The potentials (82.11) and (82.12) are called the *Lienard-Weichert potentials*, and with their aid it is possible to evaluate the electromagnetic field of an arbitrarily moving point charge.

Evaluation of the Field of an Arbitrarily Moving Point Charge. The electromagnetic field vectors are related to the potentials by the equations

$$\mathbf{E} = -\operatorname{grad}\varphi - \frac{\partial\mathbf{A}}{\partial t} \qquad \mathbf{B} = \operatorname{curl}\mathbf{A} \tag{82.13}$$

where the scalar and vector potentials are given by (82.11) and (82.12) which, for convenience, may be written

$$\mathbf{A} = \frac{\mu_0}{4\pi}\left(\frac{e\mathbf{v}}{s}\right)_\tau \qquad \varphi = \frac{1}{4\pi\epsilon_0}\left(\frac{e}{s}\right)_\tau$$
$$\tau = t - \frac{r}{c} \qquad s = r - \frac{\mathbf{r}\cdot\mathbf{v}}{c} \tag{82.14}$$

The coordinates of the motion of the electron are given as functions of time t'

$$y_1 = y_1(t') \qquad y_2 = y_2(t') \qquad y_3 = y_3(t') \tag{82.15}$$

Condition (82.1) takes the form

$$c(t - t') = r = \sqrt{[x_1 - y_1(t')]^2 + [x_2 - y_2(t')]^2 + [x_3 - y_3(t')]^2} \qquad (82.16)$$

This equation defines t' implicitly as a function of x_1, x_2, x_3, t. The vector $\mathbf{r}$, which has the components $x_i - y_i(t')$ $(i = 1, 2, 3)$, depends explicitly on x, y, z, and t' only, while the velocity vector $\mathbf{v}$, which has components $v_i = dy_i/dt'$, depends explicitly on t' only. The time t does not enter explicitly into the expressions for the potentials.

From equation (82.13), taking (82.14) into account, we obtain the following expression for the electric field $\mathbf{E}$

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\frac{e}{s^2}\operatorname{grad} s + \frac{\mu_0}{4\pi}\left(\frac{e\mathbf{v}}{s^2}\frac{\partial s}{\partial t} - \frac{e}{s}\frac{\partial \mathbf{v}}{\partial t}\right) \qquad (82.17)$$

To find the magnetic field, we use the vector analysis formula (A.14) in Appendix 1 and obtain

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{e}{s}\operatorname{curl}\mathbf{v} + \frac{\mu_0}{4\pi}e\operatorname{grad}\frac{1}{s}\times\mathbf{v} = \frac{\mu_0}{4\pi}\frac{e}{s}\operatorname{curl}\mathbf{v} - \frac{\mu_0}{4\pi}\frac{e}{s^2}\operatorname{grad} s\times\mathbf{v} \qquad (82.18)$$

Thus, the problem of finding $\mathbf{E}$ and $\mathbf{B}$ is reduced to the determination of $\partial s/\partial t$, grad s, $\partial\mathbf{v}/\partial t$ and curl $\mathbf{v}$.

Since $\mathbf{v}$ depends on x, y, z, t only implicitly, through t', while s contains x, y, z, explicitly, we obtain

$$\frac{\partial s}{\partial t} = \frac{\partial s}{\partial t'}\frac{\partial t'}{\partial t} = \left(-\frac{\mathbf{r}\cdot\mathbf{v}}{r} - \frac{v^2}{c} - \frac{\mathbf{r}\cdot\dot{\mathbf{v}}}{c}\right)\frac{\partial t'}{\partial t} \qquad (82.19)$$

taking into account that

$$\frac{\partial\mathbf{r}}{\partial t'} = \mathbf{v} \qquad \frac{\partial r}{\partial t'} = -\frac{\mathbf{r}\cdot\mathbf{v}}{r}$$

Furthermore, we have

$$\frac{\partial\mathbf{v}}{\partial t} = \frac{\partial\mathbf{v}}{\partial t'}\frac{\partial t'}{\partial t} = \dot{\mathbf{v}}\frac{\partial t'}{\partial t} \qquad (82.20)$$

$$\operatorname{curl}_x\mathbf{v} = \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} = \dot{v}_z\frac{\partial t'}{\partial y} - \dot{v}_y\frac{\partial t'}{\partial z}$$
$$= \operatorname{grad}_y t'\dot{v}_z - \operatorname{grad}_z t'\dot{v}_y = (\operatorname{grad} t'\times\dot{\mathbf{v}})_x \qquad (82.21)$$

The other components of curl $\mathbf{v}$ are of similar form, and therefore

$$\operatorname{curl}\mathbf{v} = \operatorname{grad} t'\times\dot{\mathbf{v}} \qquad (82.22)$$

The quantities $\partial t'/\partial t$ and grad t' are obtained from (82.16), by differentiating with respect to t and taking the gradient, respectively

$$c\left(1 - \frac{\partial t'}{\partial t}\right) = \frac{\partial r}{\partial t'}\frac{\partial t'}{\partial t} = -\frac{\mathbf{r}\cdot\mathbf{v}}{r}\frac{\partial t'}{\partial t}$$

$$-c \operatorname{grad} t' = \operatorname{grad} r = \frac{\mathbf{r}}{r} - \frac{\mathbf{r}\cdot\mathbf{v}}{r}\operatorname{grad} t' \tag{82.23}$$

Consequently

$$\frac{\partial t'}{\partial t} = \frac{1}{1 - \dfrac{\mathbf{r}\cdot\mathbf{v}}{rc}} = \frac{r}{s} \qquad \operatorname{grad} t' = \frac{-\mathbf{r}}{rc\left(1 - \dfrac{\mathbf{r}\cdot\mathbf{v}}{rc}\right)} = -\frac{\mathbf{r}}{cs} \tag{82.24}$$

Substituting these expressions in (82.17) and (82.18), we obtain the following expressions for $\mathbf{E}$ and $\mathbf{H}$

$$\left.\begin{aligned} 4\pi\epsilon_0\mathbf{E} &= -\frac{e}{s^2}\frac{r}{c^2}\dot{\mathbf{v}} + \frac{er}{s^3}\left(\frac{\mathbf{r}}{r} - \frac{\mathbf{v}}{c}\right)\left(1 - \frac{v^2}{c^2} + \frac{\mathbf{r}\cdot\dot{\mathbf{v}}}{c^2}\right) \\ 4\pi\mathbf{H} &= -\frac{e}{s^2}\frac{\mathbf{r}\times\dot{\mathbf{v}}}{c} + \frac{ec}{s^3}\mathbf{v}\times\mathbf{r}\left(1 - \frac{v^2}{c^2} + \frac{\mathbf{r}\cdot\dot{\mathbf{v}}}{c^2}\right) \end{aligned}\right\} \tag{82.25}$$

We can verify directly that

$$\frac{\mathbf{H}}{c} = \frac{\mathbf{r}}{r} \times \mathbf{D} \tag{82.26}$$

i.e., the magnetic field is everywhere perpendicular to the radius vector. The electric field has, however, a radial component

$$E_r = \frac{1}{4\pi\epsilon_0}\frac{e}{s^2}\left(1 - \frac{v^2}{c^2}\right) \tag{82.27}$$

The field defined by (82.25) may be resolved into two parts:

(a) the field which depends only on the velocities; this field is inversely proportional to the square of the distance, and has the form

$$\left.\begin{aligned} 4\pi\epsilon_0\mathbf{E}_1 &= \frac{er}{s^3}\left(\frac{\mathbf{r}}{r} - \frac{\mathbf{v}}{c}\right)\left(1 - \frac{v^2}{c^2}\right) \\ 4\pi\mathbf{H}_1 &= \frac{ce}{s^3}\mathbf{v}\times\mathbf{r}\left(1 - \frac{v^2}{c^2}\right) \end{aligned}\right\} \tag{82.28}$$

(b) the field which depends on the acceleration; this field is inversely proportional to the first power of the distance, and is of the form

$$\left.\begin{aligned} 4\pi\epsilon_0\mathbf{E}_2 &= -\frac{er}{c^2s^2}\dot{\mathbf{v}} + \frac{er}{c^2s^3}(\mathbf{r}\cdot\dot{\mathbf{v}})\left(\frac{\mathbf{r}}{r} - \frac{\mathbf{v}}{c}\right) \\ 4\pi\mathbf{H}_2 &= -\frac{e}{cs^2}\mathbf{r}\times\dot{\mathbf{v}} + \frac{e}{cs^3}\mathbf{r}\cdot\dot{\mathbf{v}}(\mathbf{v}\times\mathbf{r}) \end{aligned}\right\} \tag{82.29}$$

Direct verification shows that

$$\frac{\mathbf{H}_2}{c} = \frac{\mathbf{r}}{r} \times \mathbf{D}_2 \tag{82.30}$$

and that both $\mathbf{E}_2$ and $\mathbf{H}_2$ are perpendicular to the radius vector. Taking the moduli of both sides of (82.30), we obtain

$$\frac{H_2}{c} = D_2 \tag{82.31}$$

Since $1/c = \sqrt{\epsilon_0\mu_0}$, this equation may also be put in the form

$$\sqrt{\epsilon_0}\, E_2 = \sqrt{\mu_0}\, H_2 \tag{82.32}$$

This means that the vectors of the electric and magnetic fields, which depend on the acceleration, are related in the way characteristic of an electromagnetic wave. The combined electric and magnetic fields represent a spherical electromagnetic wave emitted by the electron. The first field (82.28), which depends only on the velocities, is an electrostatic field moving with the electron, but the second field (82.29), which is due to the acceleration, is a radiation field. At large distances from the electron the radiation field predominates, since it decreases more slowly with the distance than the electrostatic field.

Radiated Energy. In order to find the total energy radiated by an electron in accelerated motion, we must find the relativistic generalization of equation (38.32), which has been obtained for low velocities. Let us consider an arbitrarily moving electron. This electron radiates electromagnetic waves because of the acceleration; these waves carry energy and momentum. Let us consider a coordinate system in which the electron is at rest at the instant of radiation, i.e., its velocity is zero, but its acceleration is nonzero. From equation (38.22), we may conclude that, as a result of the acceleration, the electron loses, by radiation, an energy equal to

$$\frac{dW^{(0)}}{dt} = -\frac{e^2}{6\pi\epsilon_0 c^3}(\dot{\mathbf{v}})^2 \tag{82.33}$$

However, since the electron is at rest in this system, all directions of radiation are equivalent. Hence, the electromagnetic wave is radiated isotropically in all directions and the total momentum carried away is zero. Therefore, the change in momentum due to radiation is zero in a coordinate system in which the electron is instantaneously at rest

$$\frac{d\mathbf{G}^{(0)}}{dt} = 0 \tag{82.34}$$

We must now put equations (82.33) and (82.34) in relativistically invariant form. First, we observe that, in the coordinate system in which

the electron is at rest at a given instant, the differential of the proper time $d\tau = dt$. Let us consider a wave train radiated during time $d\tau$. By (82.33), the energy of this wave train is $dW^{(0)}$, and the momentum it carries away is $d\mathbf{G}^{(0)} = 0$. It has been shown in §81 that the energy and momentum of a finite wave train constitute a four-dimensional vector. Hence, the quantities $d\mathbf{G}^{(0)}$ and $i\,dW^{(0)}/c$ in (82.33) and (82.34) form a four-dimensional vector dI_ν with the components

$$(dI_1, dI_2, dI_3, dI_4) = \left(dG_x^{(0)}, dG_y^{(0)}, dG_z^{(0)}, i\frac{dW^{(0)}}{c}\right) \tag{83.34a}$$

Equations (82.33) and (82.34) may, therefore, be written as a single vector equation

$$\frac{dI_\nu}{d\tau} = -\frac{e^2}{6\pi\epsilon_0 c^5}(\dot{\mathbf{v}}_0)^2 u_\nu \tag{82.35}$$

where u_ν are the components of the four-dimensional velocity of the electron. In the coordinate system under consideration, we have

$$(u_1, u_2, u_3, u_4) = (0, 0, 0, ic) \tag{82.36}$$

and, therefore, equation (82.35) is identical with (82.34) and (82.33).

As indicated by equations (71.15) and (71.16), in the coordinate system in which the electron is instantaneously at rest, the components of its four-dimensional acceleration b_ν are

$$(b_1, b_2, b_3, b_4) = (\dot{v}_{0x}, \dot{v}_{0y}, \dot{v}_{0z}, 0) \tag{82.37}$$

i.e., the space components of the four-dimensional acceleration are equal to the components of the ordinary three-dimensional acceleration, and the fourth component is zero. Hence, using (82.37), we may write

$$\dot{\mathbf{v}}_0^2 = \sum_\mu (b_\mu)^2 \tag{82.38}$$

Using (82.38), we may write (82.35) in a completely relativistically invariant form

$$\frac{dI_\nu}{d\tau} = -\frac{e^2}{6\pi\epsilon_0 c^5}\left(\sum_\mu b_\mu^2\right)u_\nu \tag{82.39}$$

In the special case when the electron in a given coordinate system is instantaneously at rest, these expressions describe correctly the radiation of the electron in accordance with (82.33) and (82.34). From the relativistically invariant form of (82.38), we may conclude that it also describes the radiation in other coordinate systems in which the electron is moving at a given instant with an arbitrary velocity. Taking into account equation (71.17) for the square of the acceleration in the case of arbitrary motion of the electron, and (71.8), which expresses the components of the

four-dimensional velocity in terms of three-dimensional quantities, we obtain, using (82.39), the following expressions for the change in the energy E and momentum $\mathbf{p}$ of the electron because of radiation

$$\frac{dE}{dt} = -\frac{e^2}{6\pi\epsilon_0 c^3}\frac{\dot{\mathbf{v}}^2 - \left(\frac{\mathbf{v}}{c}\times\dot{\mathbf{v}}\right)^2}{(1-\beta^2)^3} \tag{82.40}$$

$$\frac{d\mathbf{p}}{dt} = -\frac{e^2}{6\pi\epsilon_0 c^5}\mathbf{v}\frac{\dot{\mathbf{v}}^2 - \left(\frac{\mathbf{v}}{c}\times\dot{\mathbf{v}}\right)^2}{(1-\beta^2)^3} \tag{82.41}$$

Equation (82.41) describes the reaction of the radiation on the electron, which reduces the velocity of the electron. This reaction produces a decelerating force which acts on the electron. Equation (82.40) describes the rate of loss of the energy of the electron because of radiation. From the laws of conservation of energy and momentum, we may conclude that these equations describe the radiated energy and momentum during the arbitrary motion of an arbitrary charged particle, e.g., an electron. Thus, if the law of motion of the electron is known, (82.40) can be used to determine the radiated energy for that motion.

Radiation of an Electron in a Homogeneous Magnetic Field. As an example, let us consider the radiation of an electron moving in a homogeneous magnetic field with the induction $\mathbf{B}$. The electron moves in a circle, and

$$\frac{mv^2}{r} = evB \tag{82.42}$$

The frequency of revolution of the electron is

$$\omega = \frac{v}{r} = \frac{eB}{m} \tag{82.43}$$

and the acceleration $\dot{\mathbf{v}}$ is perpendicular to the velocity $\mathbf{v}$ and its absolute value is

$$|\dot{\mathbf{v}}| = \frac{v^2}{r} \tag{82.44}$$

Consequently, we must put in (82.40)

$$\dot{\mathbf{v}}^2 = \frac{v^4}{r^2} \qquad \left(\frac{\mathbf{v}}{c}\times\dot{\mathbf{v}}\right)^2 = \frac{v^4}{r^2}\frac{v^2}{c^2} \tag{82.45}$$

Thus, we obtain the following expression for the radiated power

$$\frac{dW}{dt} = -\frac{dE}{dt} = \frac{1}{6\pi\epsilon_0}\frac{v^4e^2}{c^3r^2}\frac{1}{(1-\beta^2)^2} \tag{82.46}$$

At high electron energies, when the electron velocity is close to the

velocity of light, we may assume that $v \approx c$. Therefore, in the relativistic region, (82.46) becomes

$$\frac{dW}{dt} = \frac{1}{6\pi\epsilon_0}\frac{ce^2}{r^2}\frac{1}{(1-\beta^2)^2} \tag{82.47}$$

Thus, the power radiated increases very rapidly with increase of the electron energy, as its velocity approaches the velocity of light.

In betatrons and synchrotrons, which will be considered in detail in §88, electrons move in circles at velocities close to the velocity of light. Therefore, the power radiated by them, according to (82.47), reaches a considerable value. These energy losses of the electron must be made up by external sources. If the energy transmitted to the electron from outside does not exceed the energy radiated, then the electron energy ceases to increase, i.e., the accelerator ceases to operate. Therefore, in betatrons it is impossible to obtain electrons with energies much higher than some hundreds of MeV. In synchrotrons, there is practically no limit to the energy which can be transmitted to the electron from outside, hence, large energy losses by radiation are no obstacle to the acceleration of electrons in synchrotrons.

We must note an interesting feature of the radiation of electrons in betatrons. Although the electrons in the accelerator revolve at a definite frequency (82.43), they do not radiate at only this frequency, but at all possible multiples of that frequency. The maximum energy of radiation occurs not at the fundamental frequency ω, defined by (82.43), but at a considerably higher frequency, approximately equal to

$$\omega_{\max} \approx \omega(1-\beta^2)^{-3/2} \tag{82.48}$$

For electron energies of the order of 100 MeV, the frequency at which maximum radiation occurs exceeds the frequency of revolution by some 10 million times. This means that the wavelength of maximum radiation is about 10^7 times shorter than the circumference of the accelerator. Elementary calculations show that this wavelength is in the visible part of the spectrum. Thus, at sufficiently high energies, the radiation emitted by electrons moving in circles is visible to the eye: the electrons become "luminous." This phenomenon has been investigated experimentally in some detail. The results of the experimental observations are in good agreement with the theory.

§83. Electrodynamics of Moving Media

Equations and Tensors of the Electromagnetic Field. In the first part of the book, we have considered the phenomenological electrodynamics of

media at rest. To consider the electromagnetic phenomena of moving media, we have to formulate Maxwell's equations for them. This may be done using the requirement that the laws of nature are covariant under the Lorentz transformation. We must put the well-known Maxwell's equations for stationary media in an explicitly covariant form, i.e., in the form of tensor equations. Then, by the principle of relativity, the same form of the equations must also hold for moving media. This solves the problem of formulating Maxwell's equations for moving media. Furthermore, since the coordinate transformation formulas for tensors are known, this approach solves the problem of transforming the field vectors from one moving medium to another.

Maxwell's equation *in vacuo* are formulated in §76. However, it follows immediately from the derivation of the electromagnetic field tensors (76.3) and (76.6) and of Maxwell's equations (76.16) to (76.19), that the equations for vacuum also describe the electromagnetic field in a medium if, instead of ϵ_0 and μ_0, we write ϵ and μ. The tensor form of Maxwell's equations was obtained from Maxwell's equations for a medium at rest. But, because of the tensor nature of these equations, they hold also for moving media. Hence, the field tensors (76.3) and (76.6) and Maxwell's equations (76.16) to (76.19) also describe the electromagnetic field in an arbitrarily moving medium. The laws of transformation of the field vectors from a moving medium to a medium at rest are given by (77.6) and (77.7), if we replace ϵ_0 and μ_0 by ϵ and μ.

Four-Dimensional Current. In a medium, equation (75.8) for the four-dimensional current becomes

$$(s_1, s_2, s_3, s_4) = (j_x, j_y, j_z, ic\rho) \tag{83.1}$$

where j_x, j_y, j_z are the components of the conduction current density.

Let us consider a charge at rest in the S′ coordinate system. Then the four-dimensional current density takes the form

$$s'_1 = j'_x = 0 \qquad s'_2 = j'_y = 0 \qquad s'_3 = j'_z = 0 \qquad s'_4 = ic\rho \tag{83.2}$$

Applying the transformation formulas to go over to a coordinate system S in which the charge is moving, we obtain

$$\begin{aligned} s_1 &= j_x = \frac{\rho' v}{\sqrt{1-\beta^2}} = j_z \\ s_2 &= j_y = 0 \qquad s_3 = j_z = 0 \\ s_4 &= ic\rho = \frac{ic\rho'}{\sqrt{1-\beta^2}} \end{aligned} \tag{83.3}$$

As expected, the motion of the charge leads to the appearance of a current j_x.

Let us now consider a moving conductor carrying a current. In the S′ system, we have a conduction current but no volume charges, i.e.

$$\begin{aligned} s'_1 &= j'_x \neq 0 \qquad s'_2 = j'_y \neq 0 \\ s'_3 &= j'_z \neq 0 \qquad s'_4 = ic\rho' = 0 \end{aligned} \tag{83.4}$$

In the S coordinate system, in which the conductor is moving, we have

$$\begin{aligned} j_x &= \frac{j'_x}{\sqrt{1-\beta^2}} \\ j_y &= j'_y \qquad j_z = j'_z \\ s_4 = ic\rho &= \frac{i\beta j'_x}{\sqrt{1-\beta^2}} = i\beta j_x \end{aligned} \tag{83.5}$$

Thus, in a moving conductor carrying a current, a volume charge density

$$\rho = \frac{1}{c^2}\mathbf{v}\cdot\mathbf{j} \tag{83.6}$$

is set up, where $\mathbf{j}$ is the conduction current density in a conductor moving with a velocity $\mathbf{v}$. However, the total charge in the moving conductor is zero

$$\int \rho\, dV = \frac{\mathbf{v}}{c^2}\cdot\int \mathbf{j}\, dV = 0 \tag{83.7}$$

The appearance of a volume charge density in a moving conductor is a relativistic effect. The fact that a conductor at rest, carrying a current, is neutral is because of the fact that the densities of positive ions and electrons in the conductor are equal. However, the ions are at rest, and the electrons are moving. Let us consider some section of the conductor of length $\Delta l'_{(+)}$ containing N positive ions. Let the electrons move along the conductor with a velocity u', and let there be N electrons in a length $\Delta l''_{(-)}$ in a coordinate system in which the electrons are at rest. For the conductor to be neutral, we must have

$$\Delta l'_{(+)} = \Delta l''_{(-)}\sqrt{1-\frac{u'^2}{c^2}} \tag{83.8}$$

since $\Delta l''_{(-)}\sqrt{(1-u'^2/c^2)}$ is the length of the section containing N electrons in the coordinate system in which the conductor is at rest. If the conductor carrying a current is moving, then both the positive ions and the electrons will move, but at different velocities. If we measure, in a coordinate system at rest, the length of a section containing N positive

ions and the length of a section containing N electrons, we find that these lengths are different. Denoting the velocity of the electrons by u in the coordinate system in which the conductor moves with the velocity v, we may write

$$\Delta l_{(+)} = \Delta l'_{(+)} \sqrt{1 - \frac{v^2}{c^2}} \tag{83.9}$$

$$\Delta l_{(-)} = \Delta l''_{(-)} \sqrt{1 - \frac{u^2}{c^2}} \tag{83.10}$$

From (71.11c), we have

$$\sqrt{1 - \frac{u^2}{c^2}} = \frac{\sqrt{1 - \frac{u'^2}{c^2}}}{\sqrt{1 - \frac{v^2}{c^2}}} \left(1 - \frac{uv}{c^2}\right) \tag{83.11}$$

Hence, it is clear that

$$\Delta l_{(+)} \neq \Delta l_{(-)} \tag{83.12}$$

i.e., the charge densities of the electrons and ions in a moving conductor are not equal in magnitude. This means that a volume charge density is set up in a moving conductor.

Let us consider a moving current loop (Fig. 80). From equation (83.6),

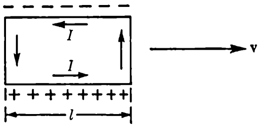


Fig. 80

the parts of the loop where the current moves in opposite directions, will carry unlike charges, as is shown in Fig. 80. Hence, the current loop acquires a dipole moment. It follows directly from Fig. 80 and equation (83.6) that this dipole moment has the magnitude

$$p = \frac{vI}{c^2} l\, l_1 = \frac{v}{c^2} IS = \frac{v}{c^2} M \tag{83.13}$$

where M is the magnetic moment of the current loop and l and l_1 are the length and width of the loop, respectively. Taking the vector nature of the terms into account, we may write this equation as

$$\mathbf{p} = \frac{1}{c^2} \mathbf{v} \times \mathbf{M} \tag{83.14}$$

Thus, the moving magnetic moment is equivalent to the electric dipole moment defined in (83.14). This fact is of great significance in atomic physics. Electrons possess magnetic moments associated with their angular momentum (spin). As they revolve about a nucleus, the electrons are in the electric field of the nucleus. The electron magnetic moment does not interact with the electric field, but, because of the effect described by (83.14), any electron moving about a nucleus must also possess an electric moment, which interacts with the electric field of the nucleus. Thus, there is an additional electrical interaction between the electron and the nucleus. This additional interaction, together with the relativistic dependence of the mass of the electron on its velocity, is responsible for the fine structure of the spectral lines of atoms.

Tensor Form of Ohm's Law. In §76 we have not derived the tensor form of the relationship called the differential form of Ohm's law

$$\mathbf{j} = \lambda \mathbf{E} \tag{83.15}$$

We shall write this relationship in tensor form in such a way that it takes the form (83.15) for a medium at rest. Direct verification shows that this tensor expression has the form

$$s_\mu = \frac{\lambda}{c} \sum_{\nu=1}^{4} H_{\mu\nu} u_\nu \tag{83.16}$$

The first three equations of (83.16) with $\mu = 1, 2, 3$ may be put in the form of a single vector equation

$$\mathbf{j} = \lambda \frac{\mathbf{E} + \mathbf{v} \times \mathbf{B}}{\sqrt{1 - \beta^2}} \tag{83.17}$$

while the fourth, with $\mu = 4$, takes the form

$$ic\rho = \frac{i\lambda}{\sqrt{1 - \beta^2}} \mathbf{E} \cdot \frac{\mathbf{v}}{c} = \frac{i}{c} \mathbf{j} \cdot \mathbf{v}$$

i.e., it is identical with equation (83.6).

Energy-Momentum Tensor of the Electromagnetic Field in a Medium. The energy-momentum tensor of the electromagnetic field in a medium may be obtained formally in a similar manner to that used in §79 for empty space. However, since $\epsilon\mu \neq 1/c^2$ for a medium, this tensor would not be symmetric; it is easy to show that the last terms of the first three rows of (79.12) are not equal to the corresponding first three terms of the last row. The lack of symmetry of the tensor gives rise to some complications. However, these can always be eliminated by using the electron theory. Study of a medium then reduces to the study of fields *in vacuo*, set up by all the charges in the medium. Hence, all problems of energy and momen-

tum may be solved in principle by means of the energy-momentum tensor for empty space (79.12). Hence, no further discussion of the tensors of the electromagnetic field in a medium is necessary.

PROBLEM

A homogeneous magnetic field H acts parallel to the axis of a dielectric cylinder of radius a and permittivity ϵ, which is rotating about its axis with an angular velocity ω. Determine the magnitude of the polarization vector of the substance of the cylinder, and the surface charge density per unit length.

Answer: $$\mathbf{P} = \frac{\epsilon - \epsilon_0}{\epsilon_0}\frac{H\omega}{c^2}\mathbf{r}$$

$$\tau = \frac{\epsilon - \epsilon_0}{\epsilon_0}\frac{H\omega}{c^2}2\pi a^2$$

CHAPTER 17

Relativistic Mechanics

§84. Equations of Motion

Newton's equations are invariant under Galilean transformations but not under the Lorentz transformation. The principle of relativity requires the laws of nature to be covariant under the Lorentz transformation. Hence, Newton's equations must be replaced by other equations of motion which are covariant under the Lorentz transformation. At low velocities, these new equations must reduce to Newton's equations.

Momentum. In classical mechanics, momentum is defined as a three-dimensional vector equal to the three-dimensional velocity vector of a particle $\mathbf{v}$ multiplied by its mass m_0. The natural relativistic generalization of this concept is the definition of a four-dimensional momentum vector G, equal to the product of the four-dimensional velocity vector of a particle u_1, u_2, u_3, u_4, and a scalar m_0 called the *rest mass* of the particle

$$(G_1, G_2, G_3, G_4) = (m_0u_1, m_0u_2, m_0u_3, m_0u_4) \tag{84.1}$$

Using equation (71.8) for the components of the four-dimensional velocity, we see that the *four-dimensional momentum* (or energy-momentum) vector may be expressed in three-dimensional terms by

$$G = \left(\frac{m_0\mathbf{v}}{\sqrt{1-\beta^2}}, \frac{im_0c}{\sqrt{1-\beta^2}}\right) \tag{84.2}$$

Why it is called the energy-momentum vector will be clear from §86.

Law of Inertia. The classical law of inertia asserts that the three-dimensional momentum of a material point, on which no external forces

are acting, is constant. The natural relativistic generalization of this law is the assertion that the four-dimensional momentum is constant in the absence of external forces

$$G = \text{const} \tag{84.3}$$

Form of the Equations of Motion. In classical mechanics, according to Newton's equations, the derivative with respect to time of the momentum of a particle is equal to the external forces acting on it. In generalizing this equation to relativistic mechanics, we must remember that it is not the ordinary time t but the proper time of the particle which is relativistically invariant. Hence, the natural generalization of Newton's second law is: the derivative with respect to the proper time of the four-dimensional momentum of a particle equals the four-dimensional force

$$\frac{dG_\nu}{d\tau} = K_\nu \tag{84.4}$$

In the case of a charged point particle in an electromagnetic field, this force is given by (78.14) and is called the Minkowski force. Using equation (78.14) for the Minkowski force, and equation (84.2) for the four-dimensional momentum, we may rewrite (84.4) in the form

$$\frac{d}{d\tau}\left(\frac{m_0\mathbf{v}}{\sqrt{1-\beta^2}}\right) = \frac{\mathbf{F}}{\sqrt{1-\beta^2}} \tag{84.5}$$

$$\frac{d}{d\tau}\left(\frac{im_0c}{\sqrt{1-\beta^2}}\right) = \frac{i}{c}\frac{e(\mathbf{v}, \mathbf{E})}{\sqrt{1-\beta^2}} \tag{84.6}$$

where $\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ is the Lorentz force. Since $d\tau = dt\sqrt{1-\beta^2}$, these equations take the form

$$\frac{d}{dt}\left(\frac{m_0\mathbf{v}}{\sqrt{1-\beta^2}}\right) = \mathbf{F} \tag{84.7}$$

$$\frac{d}{dt}\left(\frac{m_0c^2}{\sqrt{1-\beta^2}}\right) = e\mathbf{v}\cdot\mathbf{E} \tag{84.8}$$

Equations (84.7) and (84.8) are the relativistic equations of motion of a charged point particle in an electromagnetic field.

At low velocities, when $\beta \ll 1$, equation (84.7) becomes Newton's equation for the motion of a particle in an electromagnetic field

$$\frac{d}{dt}(m_0\mathbf{v}) = \mathbf{F}$$

To make the physical meaning of (84.8) clear, let us consider low velocities and assume that

$$\frac{1}{\sqrt{1-\beta^2}} = 1 + \frac{1}{2}\beta^2 + \cdots$$

Consequently, we may write

$$\frac{m_0c^2}{\sqrt{1-\beta^2}} = m_0c^2 + \frac{m_0v^2}{2} + \cdots \tag{84.9}$$

Since m_0c^2 is constant, and its derivative with respect to time equals zero, we may write (84.8), using (84.9), in the form

$$\frac{d}{dt}\left(\frac{m_0v^2}{2}\right) = e\mathbf{v}\cdot\mathbf{E} \tag{84.10}$$

Since $\mathbf{v}\,dt = d\mathbf{r}$ is the displacement of the particle, equation (84.10) means

$$d\left(\frac{m_0v^2}{2}\right) = e\,d\mathbf{r}\cdot\mathbf{E} \tag{84.11}$$

i.e., the change in the kinetic energy of a particle is equal to the work done by the electric field forces (the magnetic field does no work). Thus, equation (84.10) expresses the law of conservation of energy for the motion of a charged particle in an electromagnetic field. Hence, it is natural to suppose that in relativistic mechanics, equation (84.8) will also express the law of conservation of energy.

At present, four types of physical interaction are known: electromagnetic, gravitational, strong, and weak. The law of motion of a material point under the action of electromagnetic forces is given by equations (84.7) and (84.8). Gravitational interaction is described in the general theory of relativity, which is not discussed in this volume. The law of nuclear forces, which cause strong interactions, is still unknown. Weak interactions, which occur in the process of transformation of particles, have no effect on their macroscopic mechanical motion. Hence, equations (84.7) and (84.8) are the most important case of the relativistic equation of motion (84.4).

§85. Dependence of Mass on Velocity

Equation for the Dependence of Mass on Velocity. The mass m_0 which occurs in Newton's equation of motion is a constant for a given body and is a measure of the inertia of that body. The equation of motion in relativistic mechanics (84.7) has the form of Newton's equation, but the mass m depends on the velocity

$$\frac{d}{dt}(m\mathbf{v}) = \mathbf{F} \tag{85.1}$$

where

$$m = \frac{m_0}{\sqrt{1 - \beta^2}} \tag{85.2}$$

Thus, in relativistic mechanics the inertia of a body depends on the velocity. Equation (82.5) indicates that the mass of a body increases with increase of its velocity, and tends to infinity as the velocity tends to the velocity of light. The quantity m_0 is the value of the mass of a body at rest. It is called the rest mass of the body, and is an invariant.

Experimental Confirmation. Equation (85.2) for the dependence of the mass on the velocity has been confirmed experimentally. In the discussion of the motion of particles in transverse electric and magnetic fields in §37, it has been shown that the deflections of the particles in such fields may be used to find the ratio of the charge of a particle to its mass. The charge of a particle is constant, hence, this method may be used to determine the mass of a particle of known charge. In the derivation of formulas in §37, it has been assumed that the variation of the velocity of the particle during its motion could be ignored. This means that if we use the relativistic equation (85.1) in the derivation of these formulas, we can repeat the analysis of §37 leading to equations (37.36), (37.40), and (37.41) word for word, assuming that m is defined by (85.2) and remains constant throughout the motion. Hence, e/m must depend on the velocity. This may be checked experimentally. Bucherer's experiments, carried out in 1908 to 1909, using this method, confirmed equation (85.2) for the relationship between mass and velocity. The equation has been verified many times by more accurate methods, and each time equation (85.2) has been confirmed with great precision.

The behavior of particles in modern accelerators is a good experimental confirmation of equation (85.2). For example, in electron synchrotrons, the electrons move in closed paths in a magnetic field. The radius of curvature of the electron path in the magnetic field is found from the equation of motion, taking the projection onto the normal to the path of the particle (see eq. (37.9))

$$\frac{mv^2}{r} = evB \tag{85.3}$$

When the energy of the electron is 1000 MeV, its velocity differs from the velocity of light by about one part in ten million, and the mass is some two thousand times greater than the rest mass m_0. Experiment shows that in equation (85.3), which relates the magnetic field and the radius of curvature of the path of the electron, the mass m is, in fact, that given by

(85.2). All modern accelerator techniques are based on equation (85.2), which takes into account the dependence of mass on velocity.

Equation (85.2) has also been confirmed by many other experimental observations.

§86. Relationship Between Mass and Energy

Equation (84.8) may be written in the form

$$\frac{d}{dt}\left(\frac{m_0c^2}{\sqrt{1-\beta^2}}\right) = \mathbf{v}\cdot\mathbf{F} \tag{86.1}$$

where $\mathbf{F} = e(\mathbf{E} + (\mathbf{v} \times \mathbf{B}))$ is the Lorentz force acting on a charge in an electromagnetic field. To obtain (86.1) from (84.8), we take into account that

$$\mathbf{v}\cdot(\mathbf{v} \times \mathbf{B}) = 0 \tag{86.2}$$

since in this scalar triple product two vectors are parallel.

On the right-hand side of (86.1), we have the work done by the force $\mathbf{F}$ per unit time (since velocity = displacement per unit time). By the law of conservation of energy, the work done by an external force $\mathbf{F}$ must equal the change in the energy of the body on which the force acts. Hence

$$E = \frac{m_0c^2}{\sqrt{1-\beta^2}} = mc^2 \tag{86.3}$$

is the total energy of the body. At low velocities ($\beta \ll 1$), equation (84.9) may be put in the form

$$E = m_0c^2 + \frac{m_0v^2}{2} + \cdots \tag{86.4}$$

The quantity

$$E_0 = m_0c^2 \tag{86.5}$$

is the rest energy of the body. This must not be confused with the potential energy of a body at rest in an external field. The body possesses rest energy (86.5) in the absence of any external field; it is the internal energy of the body.

Thus, we may say that the total energy of a body consists of its rest energy (86.5) and its kinetic energy. Hence, the kinetic energy is the difference between the total energy and the rest energy

$$E_{\text{kin}} = E - m_0c^2 = m_0c^2\left(\frac{1}{\sqrt{1-\beta^2}} - 1\right) \tag{86.6}$$

At low velocities, when $\beta \ll 1$, we may use equation (86.4). Then the

kinetic energy of (86.6) reduces to the expression of classical mechanics

$$E_{\text{kin}} = \frac{m_0 v^2}{2} \tag{86.7}$$

If the velocity of a body cannot be considered to be low compared with the velocity of light, then we must use equation (86.6).

Equation (86.3), which relates the total energy E to the mass m of a body, has been deduced for the mechanical motion of a body in an electromagnetic field. But it is well known that one form of energy may be transformed into another, hence, this must be a universal relationship. The equation

$$E = mc^2 \tag{86.8}$$

is called the *mass-energy relationship*. It defines the relationship between the total energy of some body or system of bodies and the inertia of that body or system of bodies, described by the mass. It states that if a body has total energy E, then the mass of that body, which describes its inertial properties, is $m = E/c^2$. Conversely, if we know the mass of a body, we can use (86.8) to calculate its energy. Thus, (86.8) expresses a close relationship between mass and energy. In nature, energy is continuously transformed from one form to another. On the other hand, the form in which mass exists also changes. For example, an electron and a positron may collide and annihilate each other, forming a gamma-quantum in the process. The electron and positron possess rest mass, but the gamma-quantum which they form has no rest mass, although it does possess inertial mass, manifested, for example, when a gamma-quantum collides with some obstacle. Equation (86.8) states that whatever the interchange of form of the mass and energy, this mass-energy relationship will always be observed. It is incorrect to say that equation (86.8) describes the conversion of mass into energy or vice versa. We could say this only if in a given process energy disappeared and mass appeared, or vice versa. Actually, neither energy nor mass appear nor disappear; they only pass from one form into another. Mass and energy exist simultaneously, and are present in the proportions defined by (86.8).

Experimental Verification. The relationship between mass and energy (86.8) is one of the fundamental laws of modern physics, and is the basis of modern accelerator and nuclear energy techniques. The relationship has been reliably confirmed by experiment.

We can eliminate mass and velocity from equation (85.3) using (86.3), and thus, obtain a relationship between the total energy of a particle, the radius of curvature and the magnetic field. During acceleration in a synchrotron, (see §88), an electron acquires energy from the electric field

when it passes through acceleration gaps. Disregarding the so-called phase oscillations, which may be calculated, we may say that in passing through an acceleration gap the electron receives a definite amount of energy. The magnetic field must increase during each revolution of the electron by an amount necessary to keep the radius of curvature of the trajectory constant. Thus, there must exist a definite relationship between the rate of increase of the energy of the electron and the rate of increase of the magnetic field. Practical operation of accelerators shows that this relationship is the one obtained from equations (85.3) and (86.3). At the end of the acceleration, the electron energy may be measured independently, using their collisions with other particles, etc. It is thus found that the energy agrees with equation (86.3) if the velocity is expressed in terms of the magnetic field and radius of curvature by equation (85.3). During the process of acceleration, the electron energy is increased several hundredfold, and its behavior during the whole process of acceleration agrees with the theory based on equations (85.3), (86.3), and (85.2). This gives a good experimental confirmation of these equations.

Equation (86.3) also confirms the mass defect of the nucleus. It is well known that the nuclei of atoms consist of protons and neutrons, held together by nuclear forces. The laws governing nuclear forces are so far unknown. However, since these forces hold the protons and neutrons together in the nucleus, they are attractive forces. Consequently, the energy of interaction of these forces in the nucleus is negative. The total energy of the nucleus E_{nucl} consists of the rest energy of protons E_{pr} and neutrons E_{neut} and the negative energy ΔE_{int} of the interaction between protons and neutrons (this includes both potential and kinetic energy)

$$E_{\text{nucl}} = E_{\text{pr}} + E_{\text{neut}} - \Delta E_{\text{int}}$$

Hence, from (86.8), we obtain

$$m_{\text{nucl}} = m_{\text{pr}} + m_{\text{neut}} - \Delta m_{\text{int}} \tag{86.9}$$

i.e., the mass of the nucleus m_{nucl} is always less than the sum of the masses of protons m_{pr} and neutrons m_{neut} in the nucleus. The quantity Δm_{int} is called the *mass defect* of the nucleus, and serves as a measure of the forces between the nucleons in the nucleus; the greater the mass defect, the more strongly are the nucleons bound together.

The energy representing the mass defect, (86.8), is the binding energy. If the total number of nucleons is given by the mass number A, then the binding energy per nucleon ϵ equals

$$\epsilon = \frac{\Delta E_{\text{interact}}}{A} = \frac{\Delta m_{\text{interact}} c^2}{A} \tag{86.10}$$

The form of the dependence of this energy on the mass number A is shown in Fig. 81. It is clear that at the beginning of the periodic table of elements, the nucleons are bound weakly, but then the bond becomes stronger and

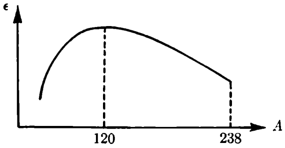


Fig. 81

reaches a maximum of approximately 8.5 MeV per nucleon, at mass numbers close to 120. The bond then begins to grow weaker, and by the end of the periodic table it has grown so weak that nuclei with mass numbers exceeding 238 do not exist in nature, and, if produced artificially, last for only a short while.

If a heavy nucleus at the end of the periodic table splits into approximately equal parts, then two nuclei of elements close to the middle of Mendeleyev's table will be obtained. According to Fig. 81, the binding energy per nucleon in these nuclei will be greater than in the original nucleus, i.e., the nucleons in these new nuclei will be more strongly bound than they were originally. Thus, the sum of the rest masses obtained by the fission of a nucleus is less than the original rest mass of the nucleus. Therefore, according to (86.8), the sum of the total rest energies after separation is less than the rest energy of the original nucleus. The difference between these energies is liberated as the kinetic energy of the fission products and as radiation. This energy is the "atomic" energy used in nuclear reactors and "atomic" bombs.

If two light nuclei at the beginning of the periodic table combine to form a single nucleus, then this nucleus will be closer to the center of Mendeleyev's table and, hence, according to Fig. 81, the nucleons will be more strongly bound than before. By a process of reasoning similar to that above, we conclude that the energy must be liberated as the result of fusion of light nuclei. This is the energy used in hydrogen bombs. The controlled use of this energy for peaceful purposes still presents great problems.

The mass of a nucleus may be determined from its motion in an electromagnetic field, by methods which are in principle the same as those used

to determine the ratio e/m. Consequently, we can also use this method to determine the mass defect of the nucleus. But, we can also measure the energies and masses of the particles involved in the fission and fusion reactions which we have just been discussing. Numerous measurements of this kind have been carried out, and these have confirmed that the relationship between mass and energy (86.8) always holds. This relationship is one of the most important equations of nuclear physics.

Relationship Between Energy and Momentum. Squaring equation (86.3), we find

$$E^2 - E^2\beta^2 = m_0^2c^4 \tag{86.11}$$

But, we may also write

$$E^2\beta^2 = m^2v^2c^2 = \mathbf{p}^2c^2 \tag{86.12}$$

where $\mathbf{p} = m\mathbf{v}$ is the momentum of a particle. Hence, we obtain the following expression for the energy

$$E = c\sqrt{p^2 + m_0^2c^2} \tag{86.13}$$

Energy-Momentum Vector of a Particle. Using equations (85.2) for the mass and (86.3) for the total energy, we may write the four-dimensional vector (84.2) in the form

$$(\mathbf{G}, G_4) = \left(\mathbf{p}, i\frac{E}{c}\right) \tag{86.14}$$

It is now clear why this vector is called the *energy-momentum vector* of the particle. The square of a four-dimensional vector is an invariant. Hence, it follows from (86.14) that

$$p^2 - \frac{E^2}{c^2} = \text{inv}$$

We shall calculate this invariant for a particle at rest, when $\mathbf{p} = 0$, $E = m_0c^2$. We obtain

$$p^2 - \frac{E^2}{c^2} = -m_0^2c^2$$

This is another form of the relationship between momentum and energy, (86.13).

Since the momentum $\mathbf{p}$ and the energy E of a particle form a single four-dimensional vector (86.14), we may write down immediately the expressions for the transformation of momentum and energy from one coordinate system to another

$$\left.\begin{aligned} p_x &= \frac{p'_x + \frac{v}{c^2}E'}{\sqrt{1-\beta^2}} \\ p_y &= p'_y \\ p_z &= p'_z \\ E &= \frac{E' + vp'_x}{\sqrt{1-\beta^2}} \end{aligned}\right\} \tag{86.15}$$

Let a particle of rest mass m_0 be at rest in the S′ coordinate system. Then we have

$$p'_x = p'_y = p'_z = 0 \qquad E' = m_0c^2 \tag{86.16}$$

Hence, from (86.15), we obtain, as expected, the following expressions for the energy and momentum

$$\begin{aligned} p_x &= \frac{m_0v}{\sqrt{1-\beta^2}} = mv \qquad p_y = p_z = 0 \\ E &= \frac{m_0c^2}{\sqrt{1-\beta^2}} = mc^2 \end{aligned} \tag{86.17}$$

Ideal Gas. As an example of the application of the mass-energy relationship, let us consider an ideal gas, assuming that the molecules collide elastically and that there are no interactions between molecules.

Let the gas as a whole be at rest in the S′ coordinate system. We denote the velocity of the i^{th} molecule in the S′ system by $\mathbf{u}'_i$. This particle has rest mass m_{0i}. We see that the energy and momentum of the gas as a whole in the S′ system are

$$E' = \sum_i \frac{m_{0i}c^2}{\sqrt{1-\frac{u'^2_i}{c^2}}} \qquad \mathbf{G}' = \sum_i \frac{m_{0i}\mathbf{u}_i}{\sqrt{1-\frac{u'^2_i}{c^2}}} \tag{86.18}$$

where the summation is performed over all the molecules of the gas. By definition of a gas at rest as a whole, $\mathbf{G}' = 0$. Applying equations (86.15) to transfer to the S coordinate system, in which the gas as a whole moves with a velocity v, we obtain

$$\begin{aligned} G_x &= \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \sum_i \frac{m_{0i}}{\sqrt{1-\frac{u'^2_i}{c^2}}} \\ E &= \frac{c^2}{\sqrt{1-\frac{v^2}{c^2}}} \sum_i \frac{m_{0i}}{\sqrt{1-\frac{u'^2_i}{c^2}}} \end{aligned} \tag{86.19}$$

Hence, it follows that, for a gas moving as a whole, the part of the rest mass M_0 is played by

$$M_0 = \sum_i \frac{m_{0i}}{\sqrt{1 - \frac{u'^2_i}{c^2}}} \tag{86.20}$$

Thus, the rest mass of the gas as a whole consists not only of the rest mass of the molecules of the gas, but also of a mass corresponding to the mass-energy relationship (86.8) of the thermal energy of the motion of the molecules. Thus, if the thermal energy of the gas increases, then the mass of the gas likewise increases. If, in addition to the kinetic energy, there is also the potential energy because of the interactions between the molecules, then the mass varies with this potential energy according to (86.8).

§87. Laws of Conservation

In classical mechanics, the momentum and energy of an isolated system are conserved. In relativistic mechanics there are similar conservation laws, but we must, of course, interpret "momentum" using the relativistic expression for momentum and "energy" as the total energy of the system. In other words, the law of conservation of energy and momentum of an isolated system is formulated in relativistic mechanics as a single law of conservation of the energy-momentum vector (86.14).

Collision of Two Particles to Form Not More Than Two Particles. From the point of view of the principle of relativity, all systems are equivalent with respect to collisions. However, this does not mean that the discussion is equally convenient in all coordinate systems. There are two coordinate systems which are especially suitable for the discussion of collisions of particles: the laboratory system of coordinates, related to the observer, and the center of mass system, related to the center of mass of the colliding particles. In the center of mass system, the particles move towards each other with velocities inversely proportional to their masses. They collide at the center of mass, and after this elastic collision, they move away from each other along lines through the center of mass, reversing their direction at the instant of collision. We must bear in mind that the direction of the line joining the centers of mass of the particles does not change its direction in space during the process of collision. The transformation of coordinates into and out of this system is effected by the transformation of the energy-momentum vector.

Let us consider the case of the collision of two particles, when not more than two particles result from the collision. The resultant particles after

the collision are, generally speaking, different from the colliding particles. A plane may be drawn through the point of collision of the colliding particles and their trajectories. By the law of conservation of momentum, the resultant particles must move in the same plane (Fig. 82.)

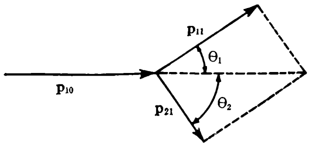


Fig. 82

We shall denote the momenta and energies before the collision by $\mathbf{p}_{10}$, $\mathbf{p}_{20}$ and E_{10}, E_{20}. The corresponding quantities after the collision are $\mathbf{p}_{11}$, $\mathbf{p}_{21}$ and E_{11}, E_{21}. We must point out once again that the resultant particles may, in general, differ from the initial particles. One of the particles (e.g., that with the index 2) may be taken to be at rest before the collision. Then the law of conservation of energy and momentum in the collision is written

$$E_{10} + m_{20}c^2 = E_{11} + E_{21} \tag{87.1}$$

$$p_{10} = p_{11} \cos \Theta_1 + p_{21} \cos \Theta_2 \tag{87.2}$$

$$0 = p_{11} \sin \Theta_1 - p_{21} \sin \Theta_2 \tag{87.3}$$

Solution of these equations enables us to find the relationship between the angles of flight and the energies of the particles after the collision.

§88. Charged Particle Accelerators

To investigate the structure of matter, the laws of interaction of different particles, and many other questions, we require high-energy particles. The natural source of such particles are the radioactive elements, e.g., radium. However, the energy of particles resulting from the radioactive decay is comparatively low, and does not exceed the 10 MeV range, while many important investigations require energies in the 1000 MeV (1 GeV) range and above. Moreover, the intensity of the particle fluxes obtained from radioactive elements is not very high.

Another natural source of high-energy particles are cosmic rays, which include particles of energies up to 10^{17}–10^{18} eV. However, the intensity of the high-energy particle fluxes in cosmic rays is very low, and, therefore,

we would have to wait a very long time to observe some effect which had low probability. For example, the very important process of the formation of a proton–antiproton pair by cosmic rays has not yet been observed, although it has long been known that this process must occur in nature.

We thus come up against the problem of accelerating particles artificially to high energies. At present, there are two methods used for particle acceleration: the *resonance method* and the *induction method.*

Principle of the Resonance Method of Acceleration. A magnetic field changes only the direction of motion of charged particles and not their velocity. Therefore, a magnetic field can be used only to control the path of a particle, and not to change its energy. Energy can be imparted to charged particles only by an electric field. If a charged particle, of charge e, equal to the electron charge, passes through an accelerating potential difference of V volts, it acquires an additional energy of eV electron volts. In practice, potential differences of several hundred thousand volts may be obtained, and, in special devices called *electrostatic generators*, we can produce potential differences of several million volts. These devices may be used to obtain particles with energies up to only several million electron volts. Given a comparatively low potential difference, the problem is, therefore, how to obtain energies many times greater than that which the particle acquires during a single passage through the given potential difference. An obvious solution of this problem is to use the same potential difference to accelerate the same particle over and over again. The simplest application of this idea consists of the following (Fig. 83). Consider a

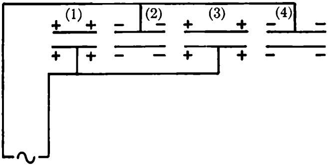


Fig. 83

succession of tubes, to which the poles of a voltage generator may be connected in turn in such a way that neighboring tubes carry unlike charges. Inside the tubes there is no field, so that a charged particle moves freely. Between the tubes an electric field is applied, which can either retard or accelerate the charged particle, depending on the sign of the

particle and the direction of the electric field between the tubes. The spaces between the tubes are called *accelerating gaps.* Suppose that at a given instant a positively charged particle enters the acceleration gap between the first and second tubes, and that the signs of the potentials are as indicated in Fig. 83. In passing through this accelerating gap, the particle gains an energy eV, where V is the potential difference. The particle then enters the second tube, where no forces act upon it. Emerging from this tube, the particle enters the accelerating gap between the second and third tubes. If the potentials were still the same, the field between the second and third tubes would decelerate the particle. To prevent this, we must reverse the potentials of all the tubes while the particle is moving through the second tube. Then, when it emerges from the second tube, the particle will once more be in an accelerating field, and will once more gain an energy eV. Thus, for the charged particle to be subjected to accelerating fields in all the accelerating gaps, the potential must be reversed every time the particle passes through a tube. For this to occur, the condition

$$\tau = \alpha \frac{T}{2} \tag{88.1}$$

must be fulfilled, where $\alpha = 1, 3, 5, \ldots$, τ is the time of flight through one tube, and T is the period of the generator. This is called the *resonance condition,* and guarantees that the particle shall move in resonance with the alternating field set up by the generator, so that it will acquire energy in passing through every accelerating gap.

During the acceleration process, the velocity of the particle increases. Hence, if the period of the generator is constant, then, to keep τ constant, the tubes must be made successively longer, so that the time of flight through each tube remains constant. It is not difficult to calculate the law of increase in length. Instead of the succession of tubes shown in Fig. 83, we could have a single accelerating gap, e.g., between the first and second tubes. We would then have to make the second tube sufficiently long and bent in, e.g., a circle, so that its end would come back to the beginning of the first tube. A magnetic field of suitable configuration would make the charged particle move inside the bent tube. By choosing the frequency of the accelerating voltage in such a way that the particle always encounters an accelerating field in the gap between the tubes, we obtain an accelerator of synchrotron or synchrocyclotron type, which we shall discuss in detail later. This method of accelerating charged particles is called the resonance method.

Induction Method of Acceleration. According to Faraday's law of electromagnetic induction, an induced electric field is set up whenever a

magnetic field changes. This induced field will accelerate charged particles. The particles move in a magnetic field in circles. Thus, when the magnetic field changes, a particle gains energy and changes its course. The question then arises as to whether or not there exist conditions under which a particle accelerated by an induced electric field would move in a circle of constant radius. Such conditions do, in fact, exist. Acceleration by means of an induced electric field is called the induction method of acceleration, and is used in the betatron.

The Cyclotron. The simplest accelerator using the resonance method is the *cyclotron*.

The frequency of revolution of a charged particle in a homogeneous magnetic field is

$$\omega = \frac{eB}{m} \tag{88.2}$$

If we can ignore the dependence of the mass of the particle on its velocity, then this frequency may be taken as constant. Equation (85.2) shows that the dependence of mass on velocity may be ignored for low velocities, when $v^2 \ll c^2$. But in the nonrelativistic case, the kinetic energy of the particle is expressed by equation (86.7). Hence, the condition $v^2 \ll c^2$ may also be written

$$\frac{v^2}{c^2} = \frac{2E_{\text{kin}}}{m_0 c^2} \ll 1$$

i.e., we may ignore the dependence of mass on velocity when the kinetic energy of the particle is much less than its rest energy. The rest energy of an electron is approximately 0.5 MeV, and the rest energy of a proton is approximately 900 MeV. Hence, at 1 MeV a proton has a low velocity and an electron has a high velocity, such that the dependence of mass on velocity plays an important role. Hence, in dealing with energies in the 10 MeV range, we may ignore the dependence of mass on velocity for heavy particles, but not for light particles. In a cyclotron, we use the constant frequency of revolution of a particle in a homogeneous magnetic field (88.2), and, hence, this apparatus is suitable for the acceleration of heavy particles but not for electrons.

A cyclotron is shown schematically in Fig. 84. A homogeneous magnetic field is applied perpendicular to the plane of the sketch, and a charged particle moves in this field at a constant frequency given by (88.2). The motion takes place through *dees*, between which we apply an alternating voltage of frequency equal to the frequency of revolution of the particle in the magnetic field. The source of the particles is near the center of the cyclotron.

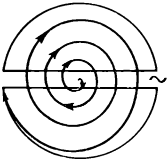


Fig. 84

The field between the dees captures the particles formed by the source, and imparts a certain amount of energy to them. When a particle moves into the space within a dee, it travels along the circumference of a circle of given radius, and, after a semiperiod of revolution, it returns to the space between the dees. The sign of the alternating voltage, and, hence, the field, has meanwhile been reversed. Therefore, the particle is once again in an accelerating field. It passes through the accelerating gap, where it receives some increase of energy, and moves into the other dee, where it moves around a semicircle somewhat greater in radius than that during the preceding half period. Then the whole process is repeated again. Thus, the particle is accelerated every time it crosses the gap between the dees, and, hence, the energy of the particle and the radius of its trajectory increase, while the period remains constant. Thus, the trajectory of the particle is a diverging spiral. The maximum energy of the particle is determined by the magnetic field and the radius of the cyclotron. When the particles attain their maximum energy, they are extracted from the cyclotron by a special device for use in further investigations.

Vertical Stability of the Particle Motion in a Cyclotron. To prevent the particles striking the horizontal walls of the dees during the acceleration process, when they deviate from the median plane of the cyclotron, we need a force which would act on them in such a way as to cause them to return to the median plane, i.e., we need to ensure the vertical stability of the motion.

There are three factors which ensure the vertical focussing of the particles in a cyclotron. First of all, they are focussed by the change in the velocity when they pass through the accelerating gap between the dees. The equipotential surfaces of the electric field between the dees are shown in Fig. 85 by the continuous lines, and the direction of the electric field vector at different points is shown by the arrows. Let us consider, for example, a positively charged particle. To be accelerated, it must pass through the gap from left to right (Fig. 85). If the particle moves strictly in the median

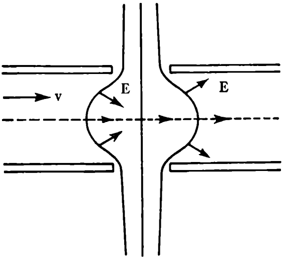


Fig. 85

plane of the cyclotron, then no other forces act on it apart from the accelerating force. But if the particle deviates from the median plane, then, as is immediately evident from Fig. 85, an additional force acts on the particle in the direction of the median plane until the particle reaches the center of the gap. After passing the center of the gap, the additional force is directed away from the medial plane. Thus, focussing takes place in the first half of the gap, and defocussing in the second half. The change in the momentum of the particle under the action of a force $\mathbf{F}$ is equal to

$$d\mathbf{p} = \mathbf{F}\,dt \tag{88.3}$$

The velocity of the particle increases on passing through the accelerating gap. Therefore, the second half of the accelerating gap is crossed in a shorter time than the first half. Hence, although the defocussing forces in the second half of the gap equal the focussing forces in the first half, the momentum imparted to the particle by the defocussing forces is less than the momentum imparted by the focussing forces. Thus, the overall effect of the forces experienced by the particle in passing through the gap is to focus the particle.

The second focussing factor is the change in the field when the particle crosses the gap. If the particle enters the accelerating gap when the electric field is increasing, the average accelerating field will be less in the first half than in the second. Hence, the focussing effect will be diminished in comparison with the defocussing effect. Consequently, if the particle crosses the gap when the field is increasing, the change in the field will have a defocussing effect. But if the particle crosses the accelerating gap when the field is decreasing, the reverse situation occurs, and the change in the

field has a focussing effect. Hence, only those particles which cross the gap when the field is decreasing will be accelerated stably. Fig. 86 shows the dependence of the voltage V on time. Let us suppose that the particle

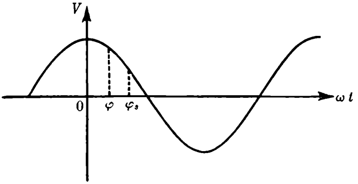


Fig. 86

is accelerated by the positive values of V. We shall take the phase origin at the point where the potential is at maximum. If the particle crosses the accelerating gap, not at the instant of maximum potential, but a little before or after it, it will acquire energy equal to

$$\Delta W = eV \cos \varphi \tag{88.4}$$

The angle φ is the transit phase of the particle in the accelerating gap. Energy may be gained during the negative phase as well as during the positive phase of equal absolute magnitude, but only the positive phase, during which the field is decreasing, will be stable. Hence, the acceleration in a cyclotron is effected during the positive phases of the alternating accelerating field.

It is easy to show that, as the energy of the particle increases, both these focussing factors become weaker. Hence, we must use additional focussing by a magnetic field. For this reason, we do not use a homogeneous magnetic field, since in a homogeneous magnetic field the lines of force are straight lines perpendicular to the median plane of motion of the particles. The Lorentz force acts toward the center of the cyclotron, parallel to the median plane. Hence, when the particle deviates from the median plane, the homogeneous field provides no forces tending to restore it to the median plane, i.e., there is no vertical magnetic focussing. To produce such a force, the lines of force must be convex, bulging outward from the center of the cyclotron (Fig. 87). Since the Lorentz force is perpendicular to the magnetic field, it is immediately clear from Fig. 87 that this force has components on both sides of the median plane which act toward this plane. Thus, this configuration of the magnetic lines of force provides vertical

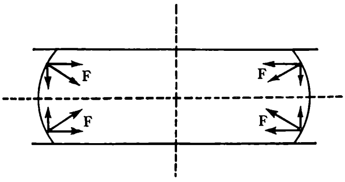


Fig. 87

focussing of the particles: when the particle deviates from the median plane, a magnetic force acts upon it, tending to restore it to the median plane. It is easy to see that the magnetic lines of force bulge in the direction of decrease of the field (this may be checked by recalling the magnetic field between the poles of a magnet). Thus, vertical magnetic focussing is ensured by a radially decreasing magnetic field.

Magnetic focussing is also used in the cyclotron. Since the magnetic field in the cyclotron decreases radially, the condition of constant frequency of revolution will not be strictly observed. Moreover, the condition of constant frequency will break down due to the increase in mass. Since the field in the accelerating gap is of constant frequency, it is impossible to accelerate the particles indefinitely; there will come a point at which the resonance breaks down and the particle crosses the gap not when the field is accelerating, but when it is decelerating, so that the acceleration of the particle stops. Thus, we may say that the limit to the energy attained by particles in a cyclotron is set by the relativistic increase of mass with velocity. Theoretically, the limit of energy is some tens of MeV; in practice it is 15 to 20 MeV.

The Betatron. The *betatron* is the only type of accelerator using the induction principle. Unlike the cyclotron, the betatron is designed for the acceleration of light particles, whose mass varies markedly with velocity.

The condition under which the accelerated electrons move in circles of constant radius is called the *betatron condition.* Consider an electron moving along a circle of constant radius r in a magnetic field increasing with time. The induced electric field is directed along the tangent to this circle of constant radius. Denoting the momentum of the electron by p, we may write down the following equation of motion

$$\frac{dp}{dt} = eE \tag{88.5}$$

where E is the induced electric field. This is a scalar equation. At every point of the circle, p and E act along the tangent. From Faraday's law of electromagnetic induction, we have

$$E = \frac{1}{2\pi r}\frac{d\Phi}{dt} \tag{88.6}$$

where Φ is the magnetic induction flux enclosed by the orbit of the electron. Substituting this expression for E in equation (88.5), and integrating with respect to time, we find

$$\int_{t_0}^{t}\frac{dp}{dt}\,dt = \frac{e}{2\pi r}\int_{t_0}^{t}\frac{d\Phi}{dt}\,dt \tag{88.7}$$

Hence, it follows that

$$p_t - p_{t_0} = \frac{e}{2\pi r}(\Phi_t - \Phi_{t_0}) \tag{88.7a}$$

We take into account the fact that

$$p = mv = eBr \tag{88.8}$$

$$\Phi = \pi r^2\langle B\rangle \tag{88.9}$$

where $\langle B\rangle$ is the mean value of the magnetic field enclosed by the electron orbit. Using (88.8) and (88.9), we find that equation (88.7) becomes

$$B_t - B_{t_0} = \frac{1}{2}(\langle B_t\rangle - \langle B_{t_0}\rangle) \tag{88.10}$$

where B is the field along the orbit, $\langle B\rangle$ is the mean field within the orbit, or simply the mean field in the betatron. Putting $B_{t_0} = \langle B_{t_0}\rangle = 0$, we obtain the betatron condition from (88.10) in the form

$$B_t = \frac{1}{2}\langle B_t\rangle \tag{88.11}$$

i.e., the field along the orbit must equal half the mean field. When this condition is satisfied, the electrons accelerated by the induced electric field move in circles of constant radius. To guarantee that (88.11) is satisfied, the field must decrease from the center of the betatron to its periphery. The distance between the poles of the magnet must, therefore, increase toward the periphery. The shape of the magnet is shown schematically in Fig. 88. The magnetic field is set up by the current flowing in the winding of the magnet. It is clear from (88.11) that the form of the dependence of the magnetic field on time plays no part at all. Hence, if the shape of the poles of the magnet is such as to satisfy (88.11), there is no need to find the dependence of the magnetic field on time.

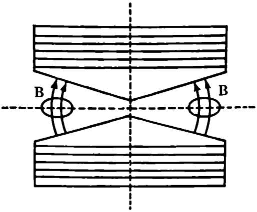


Fig. 88

The vertical stability of the motion of the electrons in a betatron is ensured by the radial decrease of the magnetic field. We still have to ensure the radial stability of the motion of the particles, i.e., we have to provide forces tending to restore the particle to a circular path of constant radius if, for some reason, it is deflected from this path. In the accelerator theory, the magnetic field is written in the form

$$B = B_0\left(\frac{r_0}{r}\right)^n = \frac{\text{const}}{r^n} \tag{88.12}$$

This is the expression for the vertical component of the magnetic field in the median plane of the accelerator. Here r_0 is the equilibrium radius of the motion, and r is the distance from the center of the accelerator to the point under discussion. To guarantee vertical stability, the condition

$$n > 0 \tag{88.13}$$

must be fulfilled. A magnetic field exerts a centripetal force on a particle moving along a circle with a given velocity v

$$F_{\text{centripetal}} = evB = evB_0\left(\frac{r_0}{r}\right)^n = \frac{\text{const}}{r^n} \tag{88.14}$$

On the other hand, the particle is subject to a centrifugal force

$$F_{\text{centrifugal}} = \frac{mv^2}{r} = \frac{\text{const}}{r} \tag{88.15}$$

Let us represent the dependence of these forces on the radius graphically (Fig. 89). At the point r_0 all three curves intersect. This corresponds to uniform motion in a circle of constant radius. Let us suppose that for some reason the particle moves out of this circle so that its distance from the center of the accelerator, for example, increases. If $n < 1$, when $r > r_0$

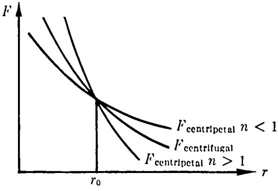


Fig. 89

the centripetal force is greater than the centrifugal, and the particle tends to return to the circle of constant radius r_0. If $n > 1$, however, the centrifugal force is greater than the centripetal, and the particle moves still further away from the position of equilibrium. Thus the motion possesses radial stability only if $n < 1$. If $n > 1$, the radial motion is unstable. This result is valid also for $r < r_0$. Taking condition (88.13) for the vertical stability into account, we may write the condition of stability of the motion of particles in a betatron in the form

$$0 < n < 1 \tag{88.16}$$

When this condition is satisfied, the particles move, during the acceleration process, along circles of constant radius, performing small oscillations, called the *betatron oscillations* about these circles.

The electrons experience acceleration as they move round the circles, and, therefore, they emit radiation. Since

$$E = \frac{m_0c^2}{\sqrt{1 - \beta^2}}$$

we may put equation (82.47) for the energy radiated by an electron moving along a circle ($v \approx c$) in the form

$$\frac{dW}{dt} = \frac{1}{6\pi\epsilon_0} \frac{ce^2}{r^2} \left(\frac{E}{m_0c^2}\right)^4 \tag{88.17}$$

Thus, as the energy of the electron increases, the power radiated increases very rapidly. The special features of this radiation were outlined in §82. The limit to the energy which can be imparted to electrons in a betatron is imposed by the energy losses due to radiation. When the electron energy reaches several MeV, the radiation becomes so strong that the induced electric field cannot compensate for these radiation losses. The betatron condition then breaks down, the electrons cease to move in circles of

constant radius, and the betatron therefore ceases to function. In practice, betatrons are designed to obtain energies of up to 300 MeV; they are not suitable for accelerating electrons to higher energies.

The Synchrotron. To obtain electrons of higher energies, we must return to the resonance principle. For energies from 4 to 5 MeV and up, the velocity of an electron is very close to the velocity of light, and we may take it to be constant, and equal to the velocity of light. Hence, if an electron moves in a circle of constant radius, its frequency of revolution will be constant. We can use a magnetic field to make the electron move in a circle and an accelerating gap to impart energy to it. Since the frequency of revolution of the electron is constant, the frequency of the field applied to the accelerating gap must also be constant. To keep the radius of the path of the electron constant as the energy increases, the magnetic field must be gradually increased. An accelerator with a variable magnetic field and a constant-frequency accelerating field is called a *synchrotron* (Fig. 90). The increase of the magnetic field and the increase of the energy

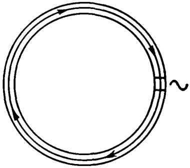


Fig. 90

of the electron are related in such a way that the electron always moves in a circle of constant radius. The radial and vertical stability is obtained in the same way as in the betatron.

Principle of Phase Stability. Let us assume that the magnetic field increases by ΔB_{rev} during one revolution. Then, during one revolution, the energy of the electron will increase by

$$\Delta E_{\text{rev}} = ecr_0\,\Delta B_{\text{rev}} \tag{88.18}$$

taking into account that $v \approx c$ and $r_0 = \text{const}$. If the amplitude of the voltage applied to the accelerating gap is V_0, then the energy increase (88.18) will be obtained on passing through the gap at a phase φ_s, which is defined by the condition (Fig. 86)

$$eV_0 \cos\varphi_s = \Delta E_{\text{rev}} \tag{88.19}$$

The phase φ_s is called the *equilibrium phase*, and the corresponding particle, the *equilibrium particle*. The equilibrium particle moves in a circle of constant radius r_0, and it crosses the accelerating gap every time at the same phase of the accelerating high-frequency field. Let us now consider the behavior of nonequilibrium particles passing through the accelerating gap at a different phase. Let us suppose that, for some reason, a certain particle crosses the accelerating gap before the equilibrium phase, i.e., $\varphi_0 < \varphi_s$, and that this particle has an energy equal to the energy of the equilibrium particle. When this particle crosses the gap, it gains more energy than does the equilibrium particle

$$eV_0 \cos \varphi_0 > eV_0 \cos \varphi_s \tag{88.20}$$

Therefore, its energy is now greater than that of the equilibrium particle. In a magnetic field

$$B = \frac{\text{const}}{r^n} \qquad 0 < n < 1$$

the radius of curvature of the path of a particle increases with increase of the particle energy. Hence, the particle under discussion will now move along a circle of greater radius than the equilibrium particle, and it will take longer to complete one revolution, since its velocity has hardly changed, remaining equal to the velocity of the equilibrium particle $v \approx c$. Thus, this particle reaches the accelerating gap next time at a phase $\varphi_1 > \varphi_0$, i.e., closer to the equilibrium phase. The particle receives, once again, more energy than the equilibrium particle, and its radius increases once more. Therefore, after the next revolution, the particle reaches the accelerating gap at a phase $\varphi_2 > \varphi_1$, and so on. Finally, after several revolutions, the particle arrives at the accelerating gap at the equilibrium phase $\varphi_n = \varphi_s$. Then the energy gain is equal to the energy gain of the equilibrium particle. But since in previous revolutions its energy gain was always greater than the gain of the equilibrium particle, the energy of the nonequilibrium particle will be greater than that of the equilibrium particle. Therefore, the nonequilibrium particle will move along a circle of larger radius, and its phase will begin to increase: $\varphi_{n+1} > \varphi_n$, $\varphi_{n+2} > \varphi_{n+1}$, etc. However, now the energy gain each time it crosses the gap will be less than the gain of the equilibrium particle. Thus, the difference in energy between the two particles will gradually decrease, becoming equal for some phase φ_{cr}. On further transits through the accelerating gap, the energy of the nonequilibrium particle now becomes less than that of the equilibrium particle energy, since it gains less energy. The radius of the nonequilibrium particle becomes, therefore, smaller than that of the equilibrium particle, and, hence, it takes a shorter time to complete one revolution. The phase

of the nonequilibrium particle begins to decrease so that it approaches again the equilibrium phase, and so on. We thus find that the phase of the nonequilibrium particles oscillates about the equilibrium phase. On the average, the nonequilibrium particles receive the same increase of energy as the equilibrium particles. Thus, if a particle is displaced from the equilibrium phase, there are factors which will tend to restore it to the equilibrium phase. This is called the *principle of phase stability*. It plays an important role in accelerators.

It follows from our description of the mechanism of phase stability that the value of the equilibrium phase is automatically selected in accordance with the rate of increase of the magnetic field: if the rate of increase of the magnetic field decreases, then the phase increases automatically; if the rate of increase of the magnetic field increases, the phase decreases. Therefore, within a wide range of conditions, it is unnecessary to hold strictly to some law of increase of the field, provided that it is not too rapid. Apart from that, the law of increase is arbitrary, since the phase of the particles will adjust itself automatically to the law of increase of the magnetic field.

There may be more than one accelerating gap in a synchrotron, enabling particles to acquire very high energy in a single revolution. The loss of energy by radiation does not, therefore, present a serious difficulty in synchrotrons, and very high energies may be obtained. There are already synchrotrons in operation in which electrons of about 6 GeV energy may be obtained. However, there is a limit to the attainable energy in a synchrotron. This limit is due to the quantum nature of radiation. The radiation of electromagnetic energy does not take place continuously, but discretely, in quanta. Under the action of this random radiation, electron oscillations appear in synchrotrons, and when these oscillations become sufficiently great, the synchrotron ceases to operate. We shall merely note that the limit to the energies which may be obtained by accelerating electrons in a synchrotron lies in the 10 GeV range.

The Synchrocyclotron. As we have shown, the cyclotron ceases to function because the frequency of revolution of the particles changes with the increase of their energy. The principle of phase stability may be used to overcome this difficulty. For this purpose, the frequency of the accelerating field must be made variable, and the relativistic change of the mass of the accelerated particles will be taken into account by the change in frequency of the accelerating field. The cyclotron with a variable frequency of the accelerating field is called the *synchrocyclotron*. The magnetic field in the synchrocyclotron is constant in time. Particles with energies of several hundred MeV may be obtained in synchrocyclotrons. In practice,

energies of about 700 MeV are obtained. Further increase of the energy is difficult for technical reasons, connected with the great weight of the magnets needed, etc. In this respect, the configuration of the field in the synchrotron is more convenient; the magnetic field is created not over the whole area of the accelerator but only in the narrow ring in which the accelerated particles move.

Proton Synchrotron. If heavy particles are accelerated in an accelerator of the synchrotron type, then, in spite of the constant radius of the path, the frequency of revolution of these particles will vary, since the velocity of heavy particles increases considerably right up to energies of 1 GeV. It is only at energies of several GeV that their velocities become close to the velocity of light, and further increase of velocity may be ignored. The frequency of the accelerating field should not, therefore, be constant, but should increase with increase of the velocity and frequency of revolution of the accelerated particles. For a particle to move in an accelerator in a circle of constant radius, the magnetic field in the ring in which the particle moves must also increase. Such an accelerator, in which the magnetic field and the frequency of the accelerating field vary but the radius of the path of the particle is constant, is called a *proton synchrotron*. It is used principally for the acceleration of protons.

The radial and vertical stability is obtained in a manner similar to the betatron. The phase stability ensures the stability of the phase oscillations, as in the synchrotron.

From the equation relating the radius of curvature of the path to the magnitude of the magnetic field, we may easily deduce the following expression (for $v \approx c$) for a field $B = 10^4$ G

$$r = \frac{10}{3} E$$

where the radius r is expressed in meters and E in GeV. It follows that an accelerator accelerating electrons to 1 GeV must have a radius of about 3.5 m. An accelerator accelerating protons to 10 GeV has a radius of about 35 m. The size of the chamber in which the acceleration takes place must be about 5% greater than the radius, so that the particles may oscillate without touching the walls. Thus, the magnetic field has to be established in large volumes, and, hence, very large magnets are needed. For example, the magnet of one particularly large accelerator, the proton synchrotron at Dubna (USSR), weighs more than 30,000 metric tons. The energy of the protons obtained from this accelerator is about 10 GeV. The weight of the magnet increases approximately as the cube of the energy. Hence, to build accelerators giving energies of 50 GeV, we would

have to make magnets weighing over a million metric tons. This is technically very difficult and very expensive.

Principle of Strong Focussing. About ten years ago a new principle of focussing particles was discovered, which has made it possible to reduce somewhat the weight of the magnets in accelerators. In ordinary accelerators, the magnetic field decreases radially from the center to the periphery according to the law (88.12), and the exponent n in this law lies between 0 and 1 (usually $n \approx 0.6$), i.e., $0 \leqslant n \leqslant 1$. The focussing forces which keep the particles in equilibrium in such accelerators are fairly weak. Hence, the amplitude of the oscillations of the particles is fairly large, i.e., the focussing of the particles is weak. Therefore, the size of the chamber in which the particles move must be fairly large, and this, in its turn, necessitates large magnets to create the necessary strong magnetic field in the chamber.

The weight of the magnets can be reduced only if the cross section of the chamber can be reduced. But this can be done only if the amplitude of the oscillations of the particles can be reduced. For this to happen, the forces which keep the particles close to the equilibrium orbit must be increased, i.e., there must be strong focussing. This can be achieved as follows. The total circumference of the accelerator is divided into a number of sectors. The magnetic field in these sectors is alternately very rapidly increasing or very rapidly decreasing. If the field is of the form (88.12), then in the sector in which the field is radially increasing, $n \ll -1$, while in the neighboring sectors, the field is rapidly decreasing radially and hence, $n \gg 1$. Usually, the absolute magnitude of n is in the range of tens or hundreds.

The theory shows that for certain relationships between the lengths of the sectors and the value of the exponent n, the motion of a particle will be stable, and the particle will be maintained in the position of equilibrium by very strong forces. For example, in an accelerator with weak focussing, a chamber of radius 1.5 m is required to attain energies of 10 GeV, but in an accelerator with strong focussing, a chamber of radius 15 cm is sufficient for energies of 30 GeV. The weight of the magnet can therefore be reduced very considerably. For example, in one of the recent strong-focussing 30 GeV accelerators, the magnet weighs 4,000 metric tons in all, while the magnet of a weak-focussing 10 GeV accelerator weighs more than 30,000 tons.

The maximum energy which can be attained at present in accelerators is about 30 GeV. A 50 to 70 GeV strong-focussing accelerator is at present under construction in the USSR.

Linear Accelerators. In all the accelerators so far described, the paths

of the particles were nearly closed curves. However, there are also accelerators in which the particles move in straight lines. These are called *linear accelerators.*

The simplest linear accelerator is a succession of tubes, to which an alternating voltage is applied, which we have already described in our discussion of the principle of resonance acceleration. However, the waveguide type of accelerator is far more widely used. If an E wave is propagated in a waveguide (see §35), then the electric field has a component along the waveguide axis. This component makes it possible to accelerate a charged particle along the waveguide axis. The velocity of the wave in the waveguide can be controlled by means of diaphragms in the waveguide. It is possible to select conditions such that the particle to be accelerated "rides" the wave, i.e., at every instant it has the same velocity as the guided wave. Therefore, an accelerating force acts on the particle all the time, and the particle gains energy.

The most powerful modern linear accelerators produce electrons with energy of about 2 GeV.

The advantage of linear electron accelerators is that the radiation is fairly weak since the magnitude of the acceleration is relatively low.

The disadvantage of linear accelerators is their great length. For example, the 30 to 40 GeV linear electron accelerator, at present under construction, will be more than 3 km long. However, linear accelerators have a number of other advantages, which we cannot discuss here, which makes the construction of such vast structures profitable.

Conclusions. Accelerator techniques are rapidly expanding at the present time. We cannot discuss every type of accelerator now in use (annular synchrocyclotron, radial-sector cyclotron, etc.), but we have confined ourselves to the description of the classical types. Most other accelerators are modifications or adaptations of those described here. However, some new types of accelerators have been recently proposed, based on different principles from those we have discussed. These topics, however, lie outside the framework of this book.

PROBLEMS

1 Find the velocity of an electron at which its mass is equal to the rest mass M_0 of a proton.

Solution:

$$\frac{m_0}{\sqrt{1-\beta^2}} = M_0,\ v = c\left[1 - \left(\frac{m_0}{M_0}\right)^2\right]^{1/2}$$

$$= c\left[1 - \frac{1}{2}\left(\frac{m_0}{M_0}\right)^2\right] = 0.999{,}999{,}843c$$

2 The rest mass of the α-particle is four times greater than the rest mass of the proton, and its charge equals two elementary charges. A proton and an α-particle are accelerated by the same potential difference. Find the potential difference for which the mass of the α-particle will be three times greater than the mass of the proton.
Solution:

$$3\left(m_0 + \frac{eV}{c^2}\right) = 4m_0 + 2\frac{eV}{c^2}$$

$$V = \frac{m_0 c^2}{e} = 9 \times 10^8\text{V}$$

3 An electron moves in a magnetic field $B = 10^4$ G, in a circle of radius 3 cm. Find the energy of the electron.
Solution: The velocity of the electron is close to the velocity of light. Hence, we may assume that $v = c$. Then we have

$$\frac{mc^2}{r} = ecB \qquad E = ecBr = 1.44 \times 10^{-10}\,J = 900\text{ MeV}$$

4 Find the increase in the mass of a kilogram of water on heating from 0 to 100° C.
Solution:

$$\Delta W = 4.18 \times 10^5\,J \qquad \Delta m = \frac{\Delta W}{c^2} = 4.7 \times 10^{-12}\text{ kg}$$

5 The solar energy flux at the earth's orbit is approximately 4 cal cm^{-2} $\text{min}^{-1} = P$. Supposing that 1% of the energy striking the earth is absorbed, determine the increase in mass of the earth per minute due to the absorption of the solar energy.
Solution:

$$\Delta W = \pi r^2 P 10^{-2} = 2.1 \times 10^{17}\text{ J/min}$$

$$\Delta m = \frac{\Delta W}{c^2} = 2.3\text{ kg/min}$$

6 What is the decrease per minute of the mass of the sun due to radiation?
Solution: Let the radius of the earth be r, and the radius of the earth's orbit be R. Then, using the result of the previous problem, we have

$$\Delta M = \frac{4\pi R^2}{\pi r^2}\Delta m 10^2 = 5 \times 10^{11}\text{ kg/min}$$

7 Find the energy liberated by 1 kg of uranium if 200 MeV are liberated on the average during the fission of a single atom of uranium. Find the equivalent quantity of coal if the caloric value of coal is $k = 7000$ kcal/kg. To what mass does this energy correspond?

Solution:

$$N\ \mathrm{kg}^{-1} = \frac{6}{235}\,10^{26} \qquad 200\ \mathrm{MeV} = 3.2 \times 10^{-11}\,J = \Delta Q$$

$$Q = \Delta Q\,N = 8 \times 10^{13}\,J = 2 \times 10^{10}\ \mathrm{k\ cal}$$

$$M_{\mathrm{equiv}} = \frac{Q}{k} = 3 \times 10^{6}\ \mathrm{kg}$$

$$m = \frac{Q}{c^2} = 0.9 \times 10^{-3}\ \mathrm{kg} = 0.9\ \mathrm{g}$$

8 How much heat (in kilocalories) is liberated during the formation of 1 kg of helium as a result of the fusion of deuterium nuclei? What is the equivalent quantity of coal, from the point of view of heat production (the caloric value of coal is 7000 kcal/kg)? How much water could the liberated heat raise from freezing point to boiling point? To what mass does this energy correspond? The binding energy of deuterium is 2.18 MeV, and of helium 28.24 MeV.
Solution: When an atom of helium is formed from two atoms of deuterium, the energy liberated is

$$\Delta Q_1 = (4 \times 28.24 - 2 \times 2.18) = 23.88\ \mathrm{MeV} = 38 \times 10^{-13}\,J$$

When one kilogram of helium is formed, the energy liberated is

$$Q = \frac{6 \cdot 10^{23}}{4}\,10^{3} \times 38 \times 10^{-13} = 5.7 \times 10^{14}\,J = 1.4 \times 10^{11}\ \mathrm{k\ cal}$$

$$M_{\mathrm{equiv}} = 2 \times 10^{7}\ \mathrm{kg} \qquad M_B = 1.4 \times 10^{9}\ \mathrm{kg} \qquad m = \frac{Q}{c^2} = 5.6\ \mathrm{g}$$

9 Calculate the time needed to accelerate a proton in a cyclotron, in which the magnetic field is $B = 10^4$ G, and the potential difference between the dees is $V = 5$ kV. The radius of the cyclotron $r_0 = 1$ m.
Solution: The maximum energy is calculated from the radius and the field

$$W_{\mathrm{max}} = \frac{e^2 B^2 r_0^2}{2m_0}$$

Hence, we find the number of revolutions of the particle during the acceleration cycle

$$N = \frac{W_{\mathrm{max}}}{2eV}$$

The period of one revolution is constant, and equals

$$T = \frac{2\pi m_0}{eB}$$

Hence, the acceleration period is

$$t = TN = \frac{\pi B r_0^2}{2V} \approx 1.57 \times 10^{-6}\ \mathrm{sec}$$

10 The radius of the orbit in a betatron is equal to $r_0 = 0.5$ m, and the angular frequency of the magnetic field is $\omega = 100\ \mathrm{sec}^{-1}$. The amplitude of the magnetic

field on the orbit is $B_0 = 5 \times 10^3$ G. Find the maximum energy which an electron can gain per revolution.
Solution:

$$\mathcal{E}_{\text{ind}} = -\frac{d\Phi}{dt} \qquad \Phi = S\langle B\rangle = 2\pi r_0^2 B_{\text{orb}}$$

$$B_{\text{orb}} = B_0 \sin \omega t \qquad \left(\frac{dB_{\text{orb}}}{dt}\right)_{\max} = \omega B_0$$

$$\Delta W = e|\mathcal{E}_{\text{ind}}| = 2\pi r_0^2 \omega B_0 e = 1.25 \times 10^{-17} J = 78 \text{ eV}$$

11 Find the energy radiated per revolution in a synchrotron, in which the orbit radius is 3.5 m, at 1 GeV energy.

$$\Delta W_{\text{rev}} = \frac{dW}{dt} T = \frac{1}{3\epsilon_0} \frac{e^2}{r} \left(\frac{E}{m_0 c^2}\right)^6 = 4.9 \times 10^{-16} J = 3000 \text{ eV}$$

12 One of the experimental confirmations of the relativistic laws of conservation of energy and momentum is provided by the Compton effect. According to quantum mechanical considerations, luminous flux may be taken to consist of individual portions of energy, light quanta. The quanta possess energy ϵ and momentum **p**, which are related to the frequency and wavelength of light by the de Broglie relationship

$$\epsilon = h\omega \qquad \mathbf{p} = h\mathbf{k}$$

$$\omega = \frac{2\pi}{T} \qquad k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$$

where λ is the wavelength of light. Here, h is Planck's constant

$$h = 1.05 \times 10^{-34} J \sec$$

In the Compton effect, when an x-ray quantum collides with a free electron, part of the energy and momentum of the quantum is transferred to the electron, and, therefore, the wavelength of the x-ray radiation scattered by free electrons increases. Find the dependence of the charge in the wavelength of the scattered radiation on the scattering angle. Assume that the electron is at rest before the collision (Fig. 91).

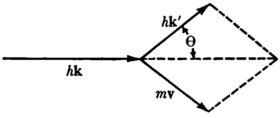


Fig. 91

Solution: Before the collision, the momentum of the quantum and its energy are equal to $h\mathbf{k}$ and $h\omega$; after the collision they are $h\mathbf{k}'$ and $h\omega'$. Before the collision, the momentum of the electron is zero, and after the collision, it is

$\mathbf{p} = m\mathbf{v}$. The energy of the electron before the collision is the rest energy m_0c^2; after the collision, it is mc^2. The laws of conservation of energy and momentum for the collision are, therefore

$$m_0c^2 + h\omega = mc^2 + h\omega'$$
$$h\mathbf{k} = h\mathbf{k}' + m\mathbf{v}$$

Rewriting these equations as

$$mc^2 = h(\omega - \omega') + m_0c^2$$
$$m\mathbf{v} = h(\mathbf{k} - \mathbf{k}')$$

and squaring, we obtain

$$m^2c^4 = h^2(\omega^2 + \omega'^2 - 2\omega\omega') + m_0^2c^6 + 2hm_0c^2(\omega - \omega')$$
$$m^2v^2 = h^2(k^2 + k'^2 - 2kk'\cos\Theta)$$

Using $k = \omega/c$, we now multiply the second equation by c, and subtract it term by term from the first equation, obtaining

$$m^2c^4\left(1 - \frac{v^2}{c^2}\right) = m_0^2c^4 - 2h^2\omega\omega'(1 - \cos\Theta) + 2hm_0c^2(\omega - \omega')$$

Taking into account that

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad 1 - \cos\Theta = 2\sin^2\frac{\Theta}{2}$$

we obtain

$$\frac{c}{\omega'} - \frac{c}{\omega} = \frac{2h}{m_0c}\sin^2\frac{\Theta}{2}$$

Since

$$\frac{c}{\omega'} = \frac{\lambda'}{2\pi} \qquad \frac{c}{\omega} = \frac{\lambda}{2\pi}$$

where λ and λ' are, respectively, the wavelengths before and after scattering, we finally obtain

$$\Delta\lambda = \lambda' - \lambda = 2\Lambda\sin^2\frac{\Theta}{2}$$

where $\Lambda = 2\pi h/m_0c = 2.42 \times 10^{-10}$ cm is the *Compton wavelength* of the electron. This is the required formula for the Compton effect, giving the change in the wavelength $\Delta\lambda$ in terms of Θ, the angle of scattering of the quantum. It is clear from this equation that the wavelength of the incident wave has no effect on the change in the wavelength, and, hence, the shorter the wavelength of the incident wave, the greater, relatively speaking, is the Compton effect. This relative change in the wavelength is especially evident for x-rays of wavelengths of the order of one angstrom (10^{-8} cm). This effect has been confirmed fully by observations, and thus the theoretical basis of the Compton equation, the relativistic laws of the conservation of energy and momentum, has also been confirmed. We recall that, according to the classical

treatment of the scattering of electromagnetic waves by free electrons (§40), the wavelength of the scattered radiation should be equal to the wavelength of the incident radiation. Therefore, the Compton effect cannot be explained in terms of the classical theory.

13 A photon of frequency ω is absorbed by an atom at rest of mass M_0. Find the velocity of the atom after absorption of the quantum.

Solution: After absorption of the photon, the atom will move in the direction of motion of the photon. Therefore, the problem is a one-dimensional one. The laws of conservation of energy and momentum take the form

$$M_0c^2 + h\omega = Mc^2$$

$$h\frac{\omega}{c} = Mv$$

From the first of these equations we obtain the mass of the atom after absorption of the photon

$$M = M_0 + \frac{h\omega}{c^2}$$

Therefore, it follows from the second equation that the velocity of the atom is

$$v = c\frac{h\omega}{M_0c^2 + h\omega}$$

Assuming that the energy of the photon is much less than the rest energy of the atom, i.e.

$$h\omega \ll M_0c^2$$

we have

$$\frac{1}{M_0c^2 + h\omega} \approx \frac{1}{M_0c^2}\left(1 - \frac{h\omega}{M_0c^2}\right)$$

Hence, the required formula in this case may be written

$$v = c\frac{h\omega}{M_0c^2}\left(1 - \frac{h\omega}{M_0c^2}\right)$$

APPENDIX 1

Vector Algebra and Analysis Formulas Used in This Book

1 Expansion of a vector triple product

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}\mathbf{A}\cdot\mathbf{C} - \mathbf{C}\mathbf{A}\cdot\mathbf{B} \tag{A.1}$$

2 Gauss' theorem

$$\int_V \operatorname{div} \mathbf{A}\, dV = \oint_S \mathbf{A}\cdot d\mathbf{S} \tag{A.2}$$

where $d\mathbf{S}$ is an element of a closed surface S enclosing a volume V. The vector $d\mathbf{S}$ points along the outward normal to S.

3 Stokes' theorem

$$\int_S \operatorname{curl} \mathbf{A}\cdot d\mathbf{S} = \oint_L \mathbf{A}\cdot d\mathbf{l} \tag{A.3}$$

where $d\mathbf{l}$ is an element of a closed contour L enclosing the surface S. The direction of describing L makes a right-hand screw system with the direction of the surface element $d\mathbf{S}$.

4 $\int_V \operatorname{curl} \mathbf{A}\, dV = \oint_S d\mathbf{S} \times \mathbf{A}$ (A.4)

where $d\mathbf{S}$ is an element of a surface S enclosing a volume V, pointing along the outward normal to S.

5 $\operatorname{div}\operatorname{grad}\varphi = \nabla^2\varphi = \Delta\varphi = \dfrac{\partial^2\varphi}{\partial x^2} + \dfrac{\partial^2\varphi}{\partial y^2} + \dfrac{\partial^2\varphi}{\partial z^2}$ (A.5)

6 $\operatorname{curl}\operatorname{grad}\varphi = 0$ (A.6)

7 $\operatorname{div}\operatorname{curl}\mathbf{A} = 0$ (A.7)

8 $\operatorname{curl}\operatorname{curl}\mathbf{A} = \operatorname{grad}\operatorname{div}\mathbf{A} - \nabla^2\mathbf{A}$ (A.8)

9 The notation

$$\mathbf{A}\cdot\nabla\mathbf{B} = A_x\frac{\partial\mathbf{B}}{\partial x} + A_y\frac{\partial\mathbf{B}}{\partial y} + A_z\frac{\partial\mathbf{B}}{\partial z} \tag{A.9}$$

10 $\operatorname{grad} \mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \nabla \mathbf{A} + \mathbf{A} \cdot \nabla \mathbf{B} + \mathbf{B} \times \operatorname{curl} \mathbf{A} + \mathbf{A} \times \operatorname{curl} \mathbf{B}$ (A.10)

11 $\operatorname{curl} \mathbf{A} \times \mathbf{B} = \mathbf{B} \cdot \nabla \mathbf{A} - \mathbf{A} \cdot \nabla \mathbf{B} + \mathbf{A} \operatorname{div} \mathbf{B} - \mathbf{B} \operatorname{div} \mathbf{A}$ (A.11)

12 $\operatorname{grad} (\varphi\psi) = \psi \operatorname{grad} \varphi + \varphi \operatorname{grad} \psi$ (A.12)

13 $\operatorname{div} (\varphi \mathbf{A}) = \varphi \operatorname{div} \mathbf{A} + \mathbf{A} \cdot \operatorname{grad} \varphi$ (A.13)

14 $\operatorname{curl} (\varphi \mathbf{A}) = \varphi \operatorname{curl} \mathbf{A} + \operatorname{grad} \varphi \times \mathbf{A}$ (A.14)

15 $\operatorname{div} \mathbf{A} \times \mathbf{B} = \mathbf{B} \cdot \operatorname{curl} \mathbf{A} - \mathbf{A} \cdot \operatorname{curl} \mathbf{B}$ (A.15)

APPENDIX 2

International (SI) System of Units

Quantity	Unit of Measurement	Dimensions	Symbol	Equivalent in Gaussian Units
		FUNDAMENTAL UNITS		
length	meter	m	m	10^2 (cm)
mass	kilogram	kg	kg	10^3 (g)
time	second	sec	sec	1 (s)
current	ampere	amp	amp	3×10^9
temperature	degree Kelvin	°K	°K	
luminous density	candle	cd	cd	
		GENERAL PHYSICAL UNITS		
velocity	meter per second	m/sec	m/sec	10^2
acceleration	meter per second per second	m/sec^2	m/sec^2	10^2
force	newton	kg m/sec^2	N	10^5 (dyne)
energy, work	joule	N m	J	10^7 (erg)
power	watt	J/sec	W	10^7
		ELECTRICAL UNITS		
quantity of electricity	coulomb	amp/sec	coul	3×10^9
voltage, potential difference, emf	volt	W/amp	V	1/300
electric field	volts per meter	V/m	V/m	$1/3 \times 10^{-4}$
electrical resistance	ohm	V/amp	Ω, ohm	$1/9 \times 10^{-11}$
capacitance	farad	coul/V	F	9×10^{11} (cm)

electrical induction	coulombs per sq. meter	coul/m^2	coul/m^2	$4\pi \times 3 \times 10^5$
		MAGNETIC UNITS		
magnetic induction flux	weber	$\frac{\text{kg} \cdot \text{m}^2}{\text{amp sec}^2} = \text{V sec}$	Wb	10^8 (maxwell)
magnetic induction	tesla	Wb/m^2	T	10^4 (gauss)
magnetic field	ampere per meter	amp/m	amp/m	$4\pi \times 10^{-3}$ (oersted)
inductance	henry	Wb/amp	H	10^9 (cm)

Selected Readings

There are a vast number of books on electrodynamics, and it would be impossible to enumerate all of them. The following is a list of some of those which the editor feels as collateral reading would make valuable contributions to the students' understanding, and which are readily available.

1. INTERMEDIATE

Corson, Dale, and Paul Lorrain, *Introduction to Electromagnetic Fields and Waves*, Freeman, San Francisco, 1962.

Feynman, R. P., R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Vol. II, Addison-Wesley, Reading, Mass., 1964.

Harwell, Gaylord P., *Principles of Electricity and Electromagnetism*, 2nd Edition, McGraw-Hill, New York, 1949.

Kip, Arthur F., *Fundamentals of Electricity and Magnetism*, McGraw-Hill, New York, 1962.

Owen, George E., *Introduction to Electromagnetic Theory*, Allyn & Bacon, Boston, 1963.

Peck, Edson R., *Electricity and Magnetism*, McGraw-Hill, New York, 1953.

Reitz, John R., and Frederick J. Milford, *Foundations of Electromagnetic Theory*, Addison-Wesley, Reading, Mass., 1960.

Schwarz, W. M., *Intermediate Electromagnetic Theory*, Wiley, New York, 1964.

Scott, William T., *The Physics of Electricity and Magnetism*, Wiley, New York, 1959.

2. ADVANCED

Abraham, M., *The Classical Theory of Electricity and Magnetism*, Hafner, New York.

Barut, A. O., *Electrodynamics and Classical Theory of Fields and Particles*, Macmillan, New York, 1964.

Becker, R., and F. Sauter, *Electromagnetic Fields and Interactions*, Vol. I, Blaisdell, New York, 1964.

Bergmann, P. G., *Introduction to the Theory of Relativity*, Prentice-Hall, Englewood Cliffs, N. J., 1942.

Jackson, J. D., *Classical Electrodynamics*, Wiley, New York, 1962.
Jeans, James, *The Mathematical Theory of Electricity and Magnetism*, 5th Edition, Cambridge, New York, 1960.
Jones, D. S., *The Theory of Electromagnetism*, Macmillan, New York, 1964.
Landau, L., and E. Lipschitz, *Electrodynamics of Continuous Media*, Addison-Wesley, Reading, Mass., 1960.
Landau, L., and E. Lipschitz, *The Classical Theory of Fields*, 2nd Edition, Addison-Wesley, Reading, Mass., 1962.
Lorentz, H. A., *The Theory of Electrons*, 2nd Edition, Dover, New York, 1952.
Moller, C., *Theory of Relativity*, Oxford U. P., New York, 1952.
Panofsky, W. K. H., and M. Phillips, *Classical Electricity and Magnetism*, 2nd Edition, Addison-Wesley, Reading, Mass., 1962.
Sommerfeld, A., *Electrodynamics*, Academic Press, New York, 1952.
Smythe, N. R., *Static and Dynamic Electricity*, 2nd Edition, McGraw-Hill, New York, 1950.
Stratton, J. A., *Electromagnetic Theory*, McGraw-Hill, New York, 1941.
Synge, John L., *Relativity: The Special Theory*, 2nd Edition, Wiley, New York, 1965.

Index